Agency : humans, animals and objects

Agency:
Humans, Animals and Objects

Thomas James Quinn

Submitted for the degree of Doctor of Philosophy

Birkbeck College, University of London

April 2017

I hereby declare that the work contained in this thesis is solely my own.

Thomas Quinn

Abstract

My aim in this thesis is to develop an account of the powers that are fundamental to human agency, by drawing out the similarities between human agency, the agency of non-human animals, and the agency of inanimate objects. Many accounts characterise our actions in terms of mental capacities unique to human agents. But focusing on what human agency has in common with agency of other kinds provides a novel perspective from which we can investigate the features of our agency that receive less attention in the literature.

I develop the account by answering two closely related questions, both of which provide the opportunity to draw out the connections between human agency, the agency of non-human animals and the agency of inanimate objects. The first question is: what are the similarities and differences between dispositions and abilities? The second question is: what are the similarities and differences between human agency and agency of other kinds?

I argue against the idea that the difference between dispositions and abilities lies in the former powers being necessitated to manifest in certain conditions. Rather, what distinguishes dispositions and abilities is that the exercise of ability involves self-movement on the part of the agent. In light of this distinction, I argue that all human actions are exercises of bodily abilities of a kind possessed by many non-human animals. Possession of these abilities does not require high-level mental capacities, but only that the agent possesses a conscious perspective. There are many ways in which the things that we do require uniquely human mental capacities, but our agency is grounded in powers of a kind held in common with non-human agents.

Acknowledgements

I am grateful to the Arts and Humanities Research Council for funding my studies.

Intellectually, my greatest debt is to my supervisor Jen Hornsby, whose guidance and uncompromising eye for detail has constantly inspired and challenged me to improve my work.

Thanks are also due to Karl Egerton, Neil Wilcock, Charlotte Knowles and Adam Ferner. I am grateful for their contributions to this thesis in various ways, but most of all for their friendship, which has been the best thing about my time at Birkbeck.

I would also like to thank Edward Wall, for reading part of this thesis and for the future book ideas; Sheffield University friends, for all the fun and the funny videos; Esther McManus, for the nature drawing; and my colleagues at the Campaign to Protect Rural England, for the afternoon teas. My parents, Colin and Ruth, and my grandma, Norma, have always supported me in whatever I have done. I should also thank my brother Ben, even though he now lives in New Zealand.

Most importantly, this thesis and much else would not have been possible without the love and encouragement of Liv Quinn.

Contents

Introduction

Bertrand Russell wrote that '[The human species] owes [its] success to certain things which distinguish [humans] from other animals: speech, fire, agriculture, writing tools, and large-scale cooperation.'[1] Whilst we could question whether Russell was right to choose these particular achievements as his examples, the point itself is certainly not a controversial one: humans have taken over the world because they can do many things that other animals cannot.

These uniquely human achievements are possible because of the unique nature of human agency. We reflect on the things that we are doing, plan for the future and engage in practical reasoning, making decisions and forming intentions based on our beliefs about the world. And these unique features of our agency are present in nearly everything we do: they are just as central to my making a cup of tea or doing the weekly shopping as to humanity's development of agriculture or taming of fire. So it is not at all surprising that many philosophers of action are most interested in these unique characteristics of human agency. My interests, however, lie elsewhere.

The aim of this thesis is to develop an account of the powers that are fundamental to human agency. But my focus is not on the characteristics that set human agency apart from agency of other kinds. I instead draw out the similarities between human agency, the agency of non-human animals, and the agency of inanimate objects. This approach provides a novel perspective from which we can investigate features of our agency that receive less attention in the literature. The focus, rather than being on the mental, is on the bodily nature of human agents' powers.

I develop this account by asking and answering two questions, both of which provide the opportunity to draw out the connections between human agency, the agency of non-human animals, and agency of inanimate objects.

The first question is: what are the similarities and differences between dispositions and abilities? The notions of disposition and ability can be used to capture the

1 (Russell 1954) p. 15

apparent contrast between human agency and the agency of inanimate objects: human action consists in the exercise of abilities, whilst the agency of inanimate objects is primarily a matter of dispositions manifesting when triggered.

The second question is: what are the similarities and differences between the agency of humans, non-human animals and inanimate objects? This question requires looking at whether these apparently separate kinds of agency involve the same types of powers. It is particularly unclear in the case of some non-human animals whether their behaviour involves the exercise of abilities, or merely the manifestation of dispositions.

These two questions are interconnected. Setting out what does and does not separate abilities and dispositions will aid in drawing lines between the different kinds of agency in which these powers play a part. And looking at the difficult cases where it is unclear whether abilities or dispositions should be attributed to an agent will require us to develop a view of what it takes to possess these different kinds of power. These questions, then, are not answered separately: they together form a thread that runs through the discussion that follows.

My discussion focuses on our concepts of agency, ability and disposition rather than the ontology of human agency and causal powers. Chapter 1 sets out to explain and justify this focus. I look at the classic debate between reductionists and antireductionists about human action, which exemplifies the disadvantages of a focus on ontology: the debate leads to a deadlock. I propose a way to make progress in the debate that should be acceptable to both sides, but which involves accepting that ontological investigation cannot tell us much of interest about the nature of human agency.

The notions of disposition and ability are introduced in Chapter 2. Dispositions are powers that are manifested in response to stimulus conditions and abilities are powers that are exercised at will. There are those who frame the difference between these two types of power as being due to the presence or lack of necessity: dispositions are necessitated to manifest by certain conditions whilst abilities are not. I examine attempts by Kadri Vihvelin and Michael Fara to analyse the notion of

ability in terms of disposition and thereby diminish this difference.[2] I conclude that this approach fails; abilities cannot be given a dispositional analysis.

Chapter 3 looks at an account of dispositions proposed by Stephen Mumford and Rani Lill Anjum, which provides a more successful way of showing that the difference between dispositions and abilities does not lie in necessity.[3] Mumford and Anjum argue that our disposition concept involves a distinctive modality weaker than necessity but stronger than possibility, which allows us to see that necessity is absent both in the case of disposition and in the case of ability. Mumford and Anjum hold that dispositions 'tend' towards their manifestations, a characteristic that does not initially seem to be possessed by abilities. I propose a view on which this 'tending' is part of a disposition's manifestation, to demonstrate that it can be seen as a feature of both dispositions and abilities.

With these possible differences between disposition and ability dismissed, Chapter 4 turns to the idea that the abilities exercised by human agents involve self-movement—a feature absent from dispositions. Through looking at accounts, proposed by Harry Frankfurt and Helen Steward, of the control over the body that is characteristic of genuine self-movement, I show that problems are prone to arise when attempting to explicate our notion of self-movement without reference to the notion of an agent exercising ability.[4]

Chapter 5 discusses a challenge to the idea that self-movement is central to human agency. Anton Ford argues that a preoccupation with self-movement has led philosophers to miss the fact that human agents directly exercise their powers on other objects in the world.[5] Many of Ford's concerns are legitimate, but by neglecting self-movement his account leaves us with no way to express the embodied nature of human agency. Filling this gap provides an opportunity to say more about the role self-movement plays in our actions. I apply the account of dispositions developed in Chapter 3 to the case of abilities, and argue that all of our interactions with other objects are exercises of bodily abilities.

2 (Vihvelin 2013); (Fara 2008)

3 (Mumford & Anjum 2011a)

4 (Frankfurt 1978); (Steward 2012)

5 (Ford 2013); (Ford 2016)

Chapter 6, the final chapter, concerns the attribution of abilities to non-human animals. I show that it is clear that some non-human animals possess bodily abilities of the kind central to human agency, but that in other cases more investigation is required. I argue that possessing these abilities requires only that an animal possess a conscious perspective, rather than requiring the high-level mental capacities distinctive of human agency. This fleshing out of the picture of bodily abilities developed in previous chapters leads to the conclusion that even animals very different to human agents possess bodily abilities of the kind central to human agency.

Chapter 1

Conceptual and Ontological Reduction

In many cases, an underlying aim of theories of human action is to show how actions are possible in the first place. It can seem as if there is a tension between our view of ourselves as agents in control of what we do and 'the world as science conceives of it'—a world of causal relations between events, or fundamental particles. In response to this tension, many philosophers of action attempt to show how actions can have the features we think they do yet still be part of a plausible naturalistic ontology.

The present chapter will show that it is possible to answer many of the most pressing questions about the nature of human agency without getting bogged down in this ontological question. Issues of ontology and issues of the content of our agential concepts can be investigated independently—and it is the latter that I will be investigating in the remainder of this thesis. The aim of this chapter, then, is to justify this approach by showing that ontological and conceptual investigation of action can come apart, and to show that favouring the ontological approach is ultimately unilluminating.

The debate between those who take a reductive approach to action and those who take an antireductive approach highlights the problem with a focus on ontological issues in the investigation of action: each side's differing views about ontology lead to a deadlock in the debate. In this chapter I will propose a new approach to investigating the nature of action that I take to be broadly antireductive. This approach also deals with the ontological question in a way that should be acceptable to both sides, and shows that, if we are interested in the nature of action, little more needs to be said about ontology. After establishing this, we will be able to move away from ontology and focus in remaining chapters on investigating our agential concepts.

1.

Reductionism and Antireductionism About Action

1.1 Reductionism

Reductionists about human action aim to resolve an apparent tension between our view of ourselves as agents and a physicist's picture of the world. Human action involves *human agents* bringing things about and exercising control over their bodies, but a scientific ontology does not include human agents. As John Bishop succinctly puts it, this tension between the two viewpoints can 'convince some philosophers either that the acting person is not part of the natural order open to scientific inquiry or that morally responsible natural agency is an illusion.'[1]

The reductionist approach to resolving this tension is to attempt to show how human action can be analysed in terms of items that clearly belong in a naturalistic ontology. To achieve this, many reductionists propose versions of the causal theory of action (CTA).[2] There have been many variations of the CTA, not all of them motivated by the desire to resolve this tension. But, broadly, the characteristic central claim of CTA accounts is that actions are bodily movements caused in the right way by mental states of the agent.[3] For example: when Anna raises her arm, she has an intention to raise her arm and this causes an event of her arm rising. Her action is the event of her arm rising, and it is an action rather than a mere bodily movement in virtue of its having the right causal history.

For those reductionists who turn to the CTA, the theory holds the promise of solving the tension between the two viewpoints because it offers a way to show that

1 (Bishop 1989) p. 15

2 E.g. (Bishop 1989, 2010); (Schlosser 2011) pp. 16-18; (Brand 1984)

3 Donald Davidson, who originated the CTA, did not appear to be motivated by this ontological question. Davidson's could even be seen as an antireductive approach, since in the face of the problem of deviant causal chains he claims that 'the concepts of event, cause and intention are inadequate to the account for intentional action'—apparently admitting that action cannot be reduced in these terms (Davidson 2004, p.106). Some have challenged this interpretation, however: Bishop (2010), who, as discussed in Section 1.3 of this chapter, maintains that ontological and conceptual reduction are separate, argues that Davidson thought that the ontological reduction of action could succeed, even if conceptual reduction could not.

the existence of human action requires nothing more than bodily movements and mental states, which seem more able to fit within a naturalistic ontology.

The reductionist does not claim to know what scientific enquiry will ultimately tell us about the nature of bodily movements or mental states. Their method is simply to take a part of our folk theory that is seen as problematic—the idea that there are human agents who engage in activity and bring things about—and then try to show how we can account for it in terms of other items in our folk theory that seem more amenable to scientific analysis. It is assumed that the mental and bodily events mentioned are the kinds of things that will eventually be accounted for in purely physical terms by scientific theory.

This route to the reductionist's proposing the CTA begins with a very restrictive assumption about what exists. But the initial tension between the agential and scientific perspective should be felt by all philosophers of action—it cannot be avoided by rejecting the assumption of a 'naturalistic ontology'. No matter one's starting point, it seems initially very plausible that the movements of our bodies during action are due to the causal work of sub-personal bodily systems—and it is initially unclear how to square this with the view that it is the *agent* who is moving the body.

Different starting assumptions about ontology simply lead to different views of what constitutes an acceptable method of resolving the tension. Those with a less restrictive view are free to make additions to the minimalist ontology favoured by many reductionists, whilst those with a more restrictive view turn to the CTA to demonstrate that, despite appearances, human agency in fact requires nothing more than what belongs to a naturalistic ontology.

To give the CTA route fair consideration, it is worth looking at why one might take on the difficult task of trying to reduce human agency instead of simply understanding human agents as beings whose actions and choices are rooted in the fact that they exist over and above the fundamental things. Some, like Bishop, argue that an account that invokes anything not mentioned by physicists gives a picture on which agency is 'mysterious' or 'supernatural'.[4] This does not provide independent

4 (Bishop 1989) p. 5

motivation for the CTA route, however, since those who have less ontologically restrictive views will simply reject the claim that anything not mentioned by physicists is supernatural.

Two other reasons for taking the CTA route are more persuasive. The first is that allowing additional things like agents into our ontology in addition to the fundamental things is not itself an easy resolution of the tension—it still requires giving an account of the relation between agents and the fundamental things. The second is that we may want to adhere to some principle of ontological parsimony, not positing additions to ontology unnecessarily—so the CTA's promise of accounting for action without proposing the existence of anything in addition to the fundamental things should be fully explored before proposing any additions to ontology.

1.2 Antireductionism

The core thesis shared by antireductive theories of action is that action is intrinsically agential and therefore cannot be reduced. One example of an antireductive view is an 'agent-causal' account of action, which was the main alternative to the CTA in the latter half of the 20th century.[5] The agent-causal view, in its classic form, tries to capture the idea that agents themselves are causally active by proposing that human actions involve a special kind of causal relation. On this view, when an agent acts there is a bodily movement caused by the agent—there is a causal relation between an *agent* and the bodily movement rather than between any bodily or mental event and the bodily movement. This introduction of a new kind of causal relation aims to afford the proper active role to the agent, rather than the causal activity being assigned to events within the agent.[6]

5 (Chisholm 1966); (Taylor 1968); (O'Connor 1999). Recent agent-causal accounts include (Alvarez 2013) and (Steward 2012).

6 Some recent agent-causal accounts argue that agents have causal powers, but deny that the agent's actions are the events brought about by exercises of these powers. (Alvarez & Hyman 1998) and (Steward 2012) identify the agent's action with the agent's exercising the power, rather than an event caused by the exercise.

Antireductive accounts need not reject the principle of ontological parsimony. The antireductionist can agree with the reductionist that we should not posit additions to ontology unnecessarily. The motivation for the antireductive position is the belief that it *is* necessary to posit additions to ontology in order to account for the existence of actions—because any attempt to reduce action is destined to fail.

One reason for thinking that reductive theories cannot succeed may come from a notorious challenge to the CTA: the problem of deviant causal chains. This is the problem of specifying in non-agential terms the exact nature of the causal connection that must obtain between mental states and bodily movements in order for somebody to be acting. This problem, more often than not, is not taken as a criticism of the CTA itself, but simply as an invitation to provide ever more complex sets of conditions in an attempt to capture the right causal connection.[7]

Here I will focus on a more overarching challenge to the CTA that provides a stronger motivation for the antireductionist. This is known as the 'disappearing agency problem'. The problem is that human agency seems to 'disappear' from an account like the CTA that mentions only bodily movements, mental events and the relations between them.

The disappearing agency problem is illustrated by A. I. Melden, who is critical of the attempt to account for action in terms of mental states causing bodily movements.[8] He writes that when mental states cause a bodily movement as on the CTA, the resulting event 'is not something *I* really willed and did, but something that was made to happen by antecedent conditions, my mental condition, my inclinations, my desires, motives and so on.'[9] The thought here is that conditions like those provided by the CTA are not sufficient for there to be an *agent* moving her body. The agent only appears in the CTA's picture as the scene of causal activity of mental states. It is these mental states, rather than the agent, that are causally active on the CTA.

Those who raise the disappearing agency problem are effectively returning to the initial tension between our conception of human agency and a restrictive naturalistic

7 (Bratman 1987) and (Mele & Moser 1994) are examples of this approach to the problem of deviant causal chains.

8 (Melden 1961)

9 Ibid. p. 7

ontology—the very problem that proponents of the CTA are trying to resolve. The problem is supposed to show that the CTA fails as a resolution of this tension, by highlighting that the tension still persists.

At this point we reach a deadlock in the debate between the reductionist and the antireductionist. The reductionist's reaction to the disappearing agency problem is simply to deny that it has any force. For example, Markus E. Schlosser writes: '[the problem's] proponents have not produced a single *argument* to support their case, and they have certainly not identified a philosophical *problem*. Their case is entirely based on intuition, and in some cases on mere metaphor and rhetoric.'[10] And it looks like the antireductionist cannot do much in the face of this outright denial that the problem needs to be answered. The two sides appear to be talking past one another.

We can begin to move past this deadlock by looking at an approach taken by John Bishop to the disappearing agency problem.[11] Bishop's aim is to defend the CTA by showing that the disappearing agency problem has no force. His interesting arguments for this conclusion, although they do not succeed, allow us to compare this debate in philosophy of action to wider debates about ontological reduction. Seeing the debate about action in this wider context will then allow us to develop an approach to action that should be satisfactory to both sides. The key to this new approach is that it moves away from focusing solely on the ontological question; it allows us to see that ontological questions are not the ones we should be most interested in when it comes to thinking about the nature of human agency.

1.3 John Bishop's response to the disappearing agency problem

Bishop initially concedes that the disappearing agency problem has some force, but argues that its reach is limited. What the problem *can* show, Bishop argues, is that our concept of action is irreducible: the notion of an agent acting is primitive and cannot be explicated in non-agential terms. This conceptual irreducibility is the reason why agency seems to disappear on the CTA picture, as the attempt to capture

10 (Schlosser 2011) p. 22

11 (Bishop 1989)

our notion of action in terms of mental states causing bodily movements fails. But, Bishop claims, the problem cannot lead to any ontological conclusions: a failure of conceptual reduction does not imply a failure of ontological reduction.[12]

If the CTA were aimed at giving a conceptual analysis of 'action', then the disappearing agency problem would straightforwardly show that the account fails. But this is not the aim of proponents of the CTA—at least, not those who turn to the CTA to resolve the conflict between our scientific and agential perspectives. They aim to show that human agency can be part of the 'natural order' by providing an ontological reduction of human actions to items in a naturalistic ontology. Bishop argues that the fact that our concept of action is irreducible still leaves it open that human agency is ontologically reducible—so the disappearing agency problem poses no threat to the CTA.

Bishop goes further than simply stating that conceptual irreducibility does not entail ontological irreducibility—he attempts to demonstrate this by showing how an ontological reduction of human agency can be given without a conceptual reduction. Rather than giving conditions that capture the meaning of 'action', Bishop suggests that in order to provide an ontological reduction the causal theorist need only give conditions that are necessary and sufficient for the occurrence of an action. These conditions would be something like: 'in all and only cases where there is an action, an agent's mental states cause a bodily movement in a certain way'. And, Bishop claims, we can give such conditions for the occurrence of something without analysing our concept of that thing.

One may raise an objection to Bishop at this point: we can often give necessary and sufficient conditions for the occurrence of something without these conditions telling us anything about the nature of that thing. For example, there could be a certain muscle in my face that moves when I smile, and which also happens to move only when I smile since it is not involved in any other of my facial movements. Here we can say that I am smiling if and only if this muscle is moving, but it is not the case that my smiling can be ontologically reduced to this muscle movement. All this shows is that it happens to be the case that this muscle moves when and only when I smile. Similarly, it may just happen to be the case that mental states cause a bodily

12 (Bishop 1989) pp. 177-180

movement in a certain way when and only when there is an action, without this telling us anything about the ontology of human agency or the nature of human actions.

Bishop is well aware of this problem, so makes his requirement for demonstrating ontological reducibility more demanding. He argues that the conditions we provide must be such that we cannot conceive of these conditions holding without there also being an action, and vice versa.[13] This, he claims, ensures that the question is not just about what happens to be the case in the actual world, but about how things would be in any possible world with the same natural ontology—and so ensures that this is a question about the nature of action. The connection between these conditions and the occurrence of an action will not be accidental, but will be due to the nature of action itself.

Bishop argues that if we can provide such conditions it is no longer reasonable to be sceptical about the possibility of reducing human agency in terms of a naturalistic ontology. Since we have shown that nothing more needs to exist than the items mentioned in the conditions in order for human agency to exist.

Bishop's defence of the CTA against the disappearing agency problem is based on a separation between conceptual and ontological reduction. He accepts that our concept of action is primitive and that the CTA's conditions cannot provide a conceptual analysis of action, but then argues that a conceptual analysis is not needed in order to show that actions can be ontologically reduced. Bishop is right to challenge the move from conceptual to ontological conclusions made by the antireductionist: it is true that conceptual irreducibility does not preclude the possibility of ontological reduction. But his defence of the CTA fails because he takes the separation between conceptual and ontological reduction too far.

Bishop's mistake is to claim both that our concept of action is primitive, and that his CTA conditions can show how action can be ontologically reduced. The former claim precludes the latter one. The fact that our concept of action is primitive may be irrelevant to ontological questions in one sense: we are not warranted in drawing conclusions about the ontology required for the existence of action on the basis of

[13] (Bishop 1989) pp. 178-179

conceptual irreducibility. But, in another sense, the conceptual irreducibility of 'action' is extremely relevant to ontological questions, in that whether or not a concept is irreducible determines how, and whether, we are able to answer ontological questions.

We can begin to understand where Bishop goes wrong by imagining how the proponent of the disappearing agency problem would respond to his method for showing that actions can be ontologically reduced. Bishop claims that we must provide necessary and sufficient conditions for the occurrence of an action, such that we cannot conceive of these conditions holding and an action failing to occur. This, he argues, is enough to show that actions are ontologically reducible. But the proponent of the disappearing agency problem will deny that it is possible to give such conditions in non-agential terms. The whole basis of the disappearing agency problem is the idea that we can conceive of the conditions specified by the CTA holding whilst there is no action taking place. The disappearing agency problem holds that when the CTA conditions are met, it is mental items that are causally active rather than the agent. Agency as we conceive of it appears to be missing from the picture where there are simply mental states causing bodily movements.

Bishop separates conceptual and ontological irreducibility, but his method for showing that action can be ontologically reduced still requires something very like a conceptual analysis of action. It requires that there is an analytic link between human agency and the conditions he provides, since this is needed for it to be the case that it is inconceivable that the conditions hold without an action taking place. This kind of analytic link is not possible if, as Bishop concedes, the disappearing agency problem shows that our concept of action is primitive. Bishop can deny that conceptual irreducibility entails ontological irreducibility, and so can maintain that the disappearing agency problem does not warrant positing additions to ontology in order to account for human agency. But he cannot, in addition to this, claim that an account of the ontological nature of action can be given based on our concept of action.

There are those who take Bishop's arguments to provide a serious defence of the CTA against the challenge of the disappearing agency problem. Aguilar and Buckareff write: 'if [Bishop's] revisionary response to the problem of the absent agent [i.e. the disappearing agency problem] is correct, drawing out ontological

conclusions that are allegedly problematic for CTA based on the concepts of 'action' and 'agency' is controversial at best and confused at worst'.[14] However, they fail to mention the problematic aspect of Bishop's argument: his further claim that he can show how actions can be ontologically reduced. Bishop may show that 'drawing out ontological conclusions that are allegedly problematic for CTA' is unwarranted, but as we have seen, he fails to show that the ontological conclusions drawn by the proponent of the CTA are any more warranted than those of the antireductionist.

Having looked at Bishop's argument and why it fails, we are now in a position to draw links between the debate about action and wider debates about ontology. The problem of fitting human agency into a 'naturalistic ontology' seems simply to be an instance of a more general problem. John Heil gives a solution to this problem that is very similar to Bishop's approach to the reduction of agency.[15] This comparison will show us how to resolve the debate between the reductionist and the antireductionist.

2.

The Debate in a Wider Context

2.1 Ontological reduction

We have seen that a motivation for some proponents of the CTA is to resolve the tension between our concept of human agency and a scientific picture of the world. We conceive of ourselves as bringing things about and exercising control over our bodies, but it is unclear how this is compatible with a scientific ontology which does not appear to include human agents. This tension is one instance of a more general conflict between our seemingly indisputable common-sense beliefs about the world we interact with every day, and the physicist's picture of what there is in the world. A famous example is found in A. S. Eddington's *The Nature of the Physical World.*

14 (Aguilar & Buckareff 2010) p. 16

15 (Heil 2003)

Eddington discovers, as he sits and writes at his desk, that there are 'two tables' before him:

> *One of them has been familiar to me from earliest years. It is a commonplace object of that environment which I call the world. How shall I describe it? It has extension; it is comparatively permanent; it is coloured; above all it is substantial. By substantial I do not merely mean that it does not collapse when I lean upon it; I mean that it is constituted of 'Substance' and by that word I am trying to convey to you some conception of its intrinsic nature[...] Table No. 2 is my scientific table. My scientific table is mostly emptiness. Sparsely scattered in that emptiness are numerous electric charges rushing about with great speed; but their combined bulk amounts to less than a billionth of the bulk of the table itself. Notwithstanding its strange construction it turns out to be an entirely efficient table. It supports my writing paper as satisfactorily as table No. 1; for when I lay the paper on it the little electric particles with their headlong speed keep on hitting the underside, so that the paper is maintained in shuttlecock fashion at a nearly steady level.*[16]

We believe that the various tables we encounter have certain features—they are solid objects with particular shapes, colours, sizes and so on. But our beliefs about what is present in these cases differ drastically from a physicist's account of what is present—in front of us are not solid objects but collections of sparsely scattered particles, or whatever the fundamental things turn out to be.[17] This conflict parallels the conflict between a common-sense conception of agents and their actions, and the physicist's picture of what exists.

The available responses to this more general conflict are also similar to the responses taken to the conflict in the case of human agency. In the case of human agency, we had the option of attempting to show that actions are reducible in terms of what is included in a naturalistic ontology, or the option of positing additions to ontology in order to account for the existence of human actions. There is a third response available, which was not considered attractive by either the reductionist or the antireductionist. This is the option of concluding that our beliefs about our own agency are mistaken, and taking the conflict to show that human agency is very unlike our conception of it.

16 (Eddington 1928) pp. ix-x

17 I leave open the question of exactly what the fundamental things are—all that matters for the purposes of this chapter is that the fundamental things are different from tables, agents, and other macro-level objects.

The responses to the more general conflict run along the same lines. We can attempt to show that tables exist by showing that they can be reduced to the fundamental things, or we can argue that objects like tables are higher-level objects that supervene on the fundamental things and are somehow realised by them. Then there is the third option: to conclude that science has shown us that, strictly speaking, tables do not exist. This third option is in fact taken by Eddington, who concludes that 'my second scientific table is the only one which is really there'.[18]

As in the case of action, none of these responses to the general conflict are without their problems. Clearly it would be strange to think that most of our beliefs about everyday objects are actually false on the grounds that objects like tables do not actually exist. If we take the approach of showing that tables are higher-level objects, much more work is required to say exactly what the relation between higher-level objects and the fundamental things is. The prospects are worse for a reductive approach: it is unclear how we can begin to give a reduction of tables in terms of fundamental things that will still allow us to accept that tables have the features we take them to possess—such as being 'constituted of "substance"'[19].

2.2 John Heil's response to the problem

I will now look at John Heil's response to this kind of problem, which is similar in a number of ways to Bishop's defence of the CTA against the disappearing agency problem.[20] Heil is concerned to avoid appealing to 'higher-level objects' or adding anything to ontology in order to account for the existence of everyday objects. Like Bishop, he suggests that progress in reconciling the tension can only be made if we separate ontological and conceptual issues. Once we do this, he argues, we can see that there is a new, more fruitful, response to the conflict. I will first look at how Heil's approach applies generally, and then at how it can help us with the issue about action.

18 (Eddington 1928) p. xii

19 Ibid. p. ix

20 (Heil 2003)

Heil identifies an assumption that he claims is the source of conflicts like the one Eddington is concerned with. He argues that many philosophers assume that by looking at our 'table' concept we can discover the nature of tables—we arrive at a set of features that anything must possess if it is to be a table. These philosophers believe that if we want to know whether tables really exist, our job is to determine whether anything in the world fulfils the criteria for being a table that we have extracted from our concept.

This methodology leads Eddington to conclude that his table does not really exist. Our concept of a table is a concept of something with a particular intrinsic nature—something 'substantial'. But science tells Eddington that he is actually sitting in front of a collection of fundamental things, which does not live up to the criteria set out for being a table. Since he believes he has discovered that nothing in the world fulfils the criteria contained in our table concept, he concludes that tables do not really exist.

The assumption that Heil has identified leads to the positing of higher-level entities. We discover that the fundamental things do not fulfil the criteria set out by our concepts of everyday objects. So, if we think that there are unacceptable consequences of saying that everyday objects like tables do not exist, and cannot see any prospect for reduction, we have to explain their existence by positing that they *do* live up to the criteria by being something over and above collections of fundamental things.

Heil argues for a different view of the role played by concepts in answering ontological questions. We should not try to extract an account of the nature of tables from our 'table' concept, and then subsequently decide whether our concept actually applies to anything in the world. Rather, we should take it to be obvious that our 'table' concept has application, and that the things we sit in front of and write on really do have the property of being solid objects with particular shapes, sizes, colours and so on. If empirical investigation tells us that there are collections of particles in the place where we think there are tables, we should not take this to conflict with the information in our table concept.[21] We have simply gained some

21 Although this is the 'scientific picture' of tables that Eddington used, it is of course unlikely that physics will tell us that particles are the fundamental things.

extra information about the nature of tables that was not to be found in our concept and so cannot be incompatible with it: tables are collections of fundamental things. Similarly, we do not find that tables are not really solid when we discover that they are collections of fundamental things—we simply find out something more about what the property of solidity is.

As well as the characteristics and properties of objects, Heil applies his line of thought to the nature of objects themselves. Once we abandon the assumption that we can 'read off' the nature of everyday objects from our concepts, we need no longer think that everyday things such as tables or people must be objects in order for our concepts to apply: 'an ontology of objects—a substance ontology—is not an ontology according to which the things we ordinarily regard as objects must turn out to be objects in a strict sense.'[22]

Imagine, for example, that it is discovered that ultimately the world is made up of one single object, and everyday things like tables and people are simply ways this object is organised in particular places. A table could be a 'thickening of the space time manifold' rather than a self-contained object or a collection of particles.[23] Heil argues that our 'table' concept would still apply, and that this discovery need not be incompatible with our everyday picture of tables as objects. Again, this is because our everyday concepts simply carve up the world in a certain way, and do not contain information about what must be the case at the level of fundamental being. Heil writes that it is only through empirical investigation that we will discover the 'deep story' about everyday objects.[24]

Heather Dyke, who develops an account very similar to Heil's, puts the same point in a different way.[25] She argues that we need to draw a distinction between the truth conditions for our claims about everyday objects and the truthmakers for those claims. As competent speakers of the language, we should know the truth conditions for the assertions we make about everyday objects: 'There is one table in this room' is true if and only if there is one solid object that is has certain features and uses. But

22 (Heil 2003) p. 177

23 Ibid. p. 190

24 Ibid. p. 11

25 (Dyke 2007)

this is compatible with competent speakers being unaware of the truthmakers for our claims. The truthmakers for my claim about the table are just whatever fundamental things empirical investigation tells us are there and to which my concept 'table' applies.

The idea central to both Heil's and Dyke's views is that the content of our concepts of everyday objects is macro-level only, and does not specify the underlying ontological nature of these objects. This means that we cannot extract from our concepts any information about the ontology required in order for them to apply. The other side of this is that no empirical findings at the micro-level about what there is in the world can conflict with our everyday concepts or show us that everyday objects like tables do not really exist.[26] This picture of the way our concepts work is plausible if we think about how our concepts are formed.

Every day we interact with objects like tables—we place things on them, bump into them, polish them—and the features of tables that we can perceive and that we care about—their shape, solidity, weight—are the macro-level features that are relevant to these interactions.[27] It is plausible that our concepts of everyday objects are formed on the basis of these interactions and perceptions to enable us to recognise and communicate about these relevant features. It makes sense, then, that our concepts would not specify anything about the 'deep story' of tables—about the ontology required in order for our concepts to have application. They are only formed to deal with our macro-level engagement with objects, so are not relevant at the level of fundamental things, which are discovered by investigation with scientific instruments or postulated on the basis of physicist's theories. Those who think that there is a conflict between our conception of tables as solid objects and their being made of sparsely connected arrangements of particles are therefore subject to a confusion. Tables' solidity only makes sense at the macro-level, and so it cannot be that by looking at the micro-level we find that tables are not solid.

---

26 On Heil's view, we can still find out through empirical evidence that individual objects are not what we thought they were. For example, I can find that a diamond I have bought is actually a fake, when I take it to a lab to be analysed and find it has a different chemical composition. But what we could not discover through empirical investigation is that diamonds do not exist. Even if we found that all diamonds have different composition to that which they were previously thought to have, we would simply be finding out something new about the nature of diamonds.

27 (Heil 2005) p. 502

With their view of the role our concepts play in ontological investigation, Heil and Dyke are able to avoid the unattractive responses to the conflict that were discussed above. Instead, they take an approach that is reductive in a sense, but that succeeds on the basis of clarifying the role our concepts can play in such an investigation. They simply claim that tables are collections of fundamental things—but that this does not mean tables do not have the characteristics (solidity, extension etc.) that we assume they do in our everyday thought.

There are many similarities between Heil's and Bishop's responses to the conflict between our common-sense view of the world and an ontology informed by physics. Both want to avoid becoming eliminativists or adding items to ontology, and both attempt to do so by showing that there *can* be ontological reduction in the absence of conceptual reduction. They argue that it is a mistake to think that conceptual analysis can lead to ontological conclusions.

As we have seen, Bishop's account of how we can demonstrate ontological reduction fails. It requires an analytic route from our concept to a set of conditions for the occurrence of an action, which is inconsistent with the idea that a competent possessor of the 'action' concept should not be expected to see how human agency fits into a naturalistic ontology. Heil does not fall into this trap—on his view, the ontological nature of everyday things cannot be discovered from the armchair. Only empirical investigation will tell us the 'deep story' about these objects.

Heil's view is controversial, relying on a rejection of a commonly accepted view of the role our concepts play in ontological investigation. However, his view provides a plausible way out of the dilemma presented by ontological reduction, and shows us a way to resolve the debate about the ontology of action. I will therefore spend some time defending his view against possible criticisms, to show that despite including controversial claims this view is still a live option.

One worry may stem from the impression that Heil believes everyday objects like tables are identical to collections of fundamental things. This identity claim seems implausible, since tables and collections of particles or fundamental things have different persistence conditions. It is possible for a table to lose some of its parts and survive, but a collection of particles cannot—further, a collection of particles can drastically change shape and survive but a table cannot. Fortunately, Heil does

not make this identity claim. He is content to use phrases like 'the table is a portion of matter' but he makes clear that this is not the 'is' of identity.[28]

Realism about tables, Heil argues, requires only that our table concept has application in the world. We have seen how he thinks our concepts can have application no matter what the physicist's account of reality ends up being, and that our sortal concepts can have application even if there are no clearly delineated objects for them to attach to. For Heil, the table 'is' whatever the physicist says there is simply in the sense that our table concept applies, at a particular time, to that particular portion of reality.

This avoids the problem with identity: our table concept can apply to a particular portion of reality even though the table may cease to exist while the portion of reality continues to exist—say, if the table is cut into small pieces. In Heil's terms, what happens in this case is just that our table concept stops applying.

A second worry is that on Heil's view our everyday talk does not come out as *genuinely true*. It may seem that if we accept the view, when we say 'there is a table here or 'there is one object here' then what we are saying is actually false. We may be able to speak as if there were tables, but our talk would simply be metaphorical. However, this is not the case, and this objection rests on a misunderstanding of Heil's argument. Heil's claim is that our everyday concepts do not carry ontological commitments, and so being realists about tables is compatible with identifying tables with collections of fundamental things. Our talk about everyday objects is still true.

It is now time to take a preliminary look at how we can use Heil's and Bishop's views to find a response to the tension in the case of human agency that will be acceptable to those on both sides of the debate. The account we end up with on this first attempt will not be entirely acceptable—the proponent of the disappearing agency problem may still object that it deflates the role of agents in their actions. I will suggest some refinements to the account that should demonstrate that it allows us to be realists about human agency and the active role human agents play in their actions.

28 (Heil 2003) p.188

3.

A Way Out of the Deadlock in the Debate

3.1 Applying Heil's account to the case of action

First, let us remind ourselves of the tension we are attempting to deal with. There are two conflicting pictures of human agents and agency: on our common-sense view human agents move their bodies and actively bring things about, but on a scientific ontology there are no human agents, only whatever the fundamental things turn out to be. It is difficult to see how an agent can be moving her body if what we think is human action is just a matter of causation between fundamental things. The agent seems simply to be the site of this fundamental-level causation rather than a genuine participant in what we think of as her actions. Antireductionists respond to this by claiming that human agents and their causal powers exist over and above the fundamental things, while the proponent of CTA argues that we can reduce human actions to causation between non-agential entities.

Using an approach like Heil's, it seems that we can resolve the tension and retain genuine human agency without proposing any additions to ontology. His discussion of the relation between concepts and reality shows us why the existence of agents capable of acting is not challenged by the empirical findings of physics. Our concept of human agency does not have any ontological implications—it does not specify anything about the ontological nature of agents or their abilities—and so it cannot conflict with empirical findings that tell us there are really only collections of fundamental things. If empirical investigation tells us that there are only collections of fundamental things, this does not show that there are no human agents or that they cannot intentionally move their bodies. It simply tells us the 'deep story' about the ontological nature of human agents and their abilities. We find that what our agential concepts apply to are actually fundamental things and their causal activities.

The idea here, as with Bishop's view, is that we can provide an ontological reduction of human agency without giving a conceptual reduction. We are unable to see how human agency might fit into the picture given by physicists because our concept is

irreducibly agential—but this does not mean human agency cannot be ontologically reduced. It is not through examining our concepts but through empirical investigation that we can find the ontological nature of human agency.

The proponents of the disappearing agency problem would perhaps be unconvinced that this approach provides a genuine vindication of our everyday thought about human agency. The idea that, through empirical investigation, we can give an account of the true ontological nature of human agents in terms of the fundamental things may still seem to be eliminativist about human agency. If the accusation of eliminativism is based on an outright failure to reconsider the view of the relation between concepts and reality that Heil argues against, then it is unlikely that anything can be said to make this approach look more appealing. But the modifications that I will make to the approach should help those who think there is a grain of truth in what Heil says yet still find it difficult to agree that this approach gives us a picture of genuine human agency. The approach we have just looked at sticks very closely to Heil's view about ontological reduction of everyday objects, but we will see that a plausible approach requires moving away from some aspects of his view.

3.2 The 'deep story' about everyday things?

Heil's approach has its merits, but goes too far in claims about what we can discover through empirical investigation. Heil claims that we can discover the 'deep story' about everyday things through empirical investigation rather than through conceptual reduction.

I agree with Heil that it is empirical investigation, rather than the examination of our concepts, that will tell us about ontology. I also agree that empirical investigation will allow us to say that nothing more is needed apart from the fundamental things in order for everyday objects to exist. But it is not the case that this tells us anything very interesting about everyday objects or gives us some insight into their deep nature that is more important than our everyday understanding of them. It is particularly important to make this point in order to mitigate any worries that accepting Heil's approach means a return to straightforward reductionism about human agency.

To see why Heil's approach does not give us a 'deep story' about everyday things, we need to get clear on how ontological reduction without conceptual reduction might be carried out. How does empirical investigation lead to the conclusion that tables are collections of fundamental things?

One thing is clear—it is not through empirical investigation of tables that it is discovered that tables are in fact collections of fundamental things. There is no need to carry out any empirical investigation into a particular object to arrive at this conclusion. Rather, the conclusion is derivative from a much more general claim: that all objects are in fact collections of fundamental things.

The discovery of what the fundamental things are comes from experiments carried out in highly controlled environments, designed to test theories about what must exist. Tests carried out in the Large Hadron Collider could not be more removed from an investigation into individual objects, and discoveries about what the fundamental things are go towards creating highly general theories. So, the claim that tables are identical to collections of fundamental things is in some sense not really a claim about tables. It is simply an application of a much more general theory that applies to all objects.

This shows why it is misleading to say that we discover a 'deep story' about tables through empirical investigation into the nature of the fundamental things. We discover something about our ontology—namely what the fundamental things are and that nothing more than these things exist. But we do not discover anything specific about tables, or any other everyday object. For questions about what tables are, the features they have and so on we must still turn to our concept and our everyday interactions with these objects.

The insight provided by Heil's view is that these at first seemingly contradictory claims can both be true. It shows us how it is that our questions about the nature of tables are not answered by saying that they are collections of fundamental things—but also shows us why it is still right to claim that tables are collections of fundamental things, rather than going down the route of positing higher-level objects. His view shows us that the answers to ontological questions cannot contradict the answers we provide to everyday objects based on our concepts: we cannot read off ontology from our concepts.

3.3 The nature of human action

With a better understanding of the limitations on what we can discover through empirical investigation into ontology, we can now take another look at the picture of action Heil's approach gives us.

The discussion in the previous section shows why, by taking this approach, we are not being reductive about action. Rather, Heil's approach shows how ontology and our conceptual understanding of agency come apart, without diminishing the importance of either one.

The ontological question about what must exist in order for actions to exist is answered through empirical investigation. But, as we have seen, the conclusions drawn through empirical investigation are highly general—they are really conclusions about what must exist in order for *anything* to exist rather than conclusions about particular objects. Further, no ontological conclusion we draw can contradict or inform our understanding of human action, because ontology is not specified by our concepts of everyday things.

For these reasons, rather than through empirical investigation it is through looking at our concepts and interrogating or everyday understanding of action that we will be able to form a theory of action and agency. Questions like 'What is it for an agent to be in control of her body?', 'How should we individuate actions?' or 'What are the prerequisites for agency?' will all be answered in this way. And, importantly, given the apparent irreducibility of our agential concepts, it still looks like the correct theory will be one in which action is an irreducibly agential phenomenon.

Conclusion

My aim in this chapter has been to justify the approach to asking and answering questions about agency that I will be taking in the remainder of this thesis. Although I will adopt what I take to be an antireductive view of agency, I will not be concerned with ontological questions about how human agency as I present it fits into a scientific ontology. My focus in the remainder of the thesis will instead be on our agential concepts.

Chapter 2

Dispositional Accounts of Ability

Having justified turning away from ontological questions about human agency, we may now turn to the questions that this thesis will seek to answer. My aim is to better understand the powers fundamental to human agency as we conceive of it, by placing human agency on a continuum with the agency of inanimate objects and of non-human animals. Examining the way we draw lines between these apparently different types of agency will allow us to see our concept of human agency in a wider context, and to focus on the features that are at the core of human agency as we conceive of it.

Of course, there are many different ways of drawing comparisons between human agency and the agency of inanimate objects and of non-human animals. Human agency is unique in many respects—we have capacities for self-reflection, planning and reasoning that are not open to many non-human animals, and certainly do not play a part in the agency of inanimate objects. Since many philosophers are most interested in these unique features of human agency, many theories of action end up emphasising what separates human agency from agency of other kinds. Focussing on the similarities between human agency and agency of other kinds provides us with a novel perspective from which we can investigate features of our agency that receive less attention in the literature.

We can begin to compare human agency with other, seemingly very different, types of agency by looking at the concepts of ability and disposition. Dispositions—powers whose manifestations are triggered in specific conditions—can be associated with the agency of inanimate objects. Abilities—powers that are exercised at will—can be associated with the agency of humans.

In this and the next chapter I will focus on the apparent contrast between the manifestations of dispositions and the exercises of abilities. The difference between the two is sometimes understood in terms of dispositions being necessitated to

manifest by certain conditions and abilities not being necessitated to manifest in any conditions. This is taken by some to show that our concepts of disposition and ability are concepts of powers of fundamentally different kinds.

I will argue that this is not the case: the differences between ability and disposition, and between the agency of inanimate objects and of humans, lie elsewhere. In this chapter I will look at an attempt to show this that is ultimately unsuccessful. In the next chapter I will look at a more successful approach to showing that the difference between our concepts of disposition and ability does not lie in necessity.

1.

Abilities and Dispositions

1.1 Dispositions

The first step is to say more about the notions of disposition and ability, and the part they play in the agency of inanimate objects and of humans.

Let us begin by looking at the notion of disposition. To attribute a disposition to an object is to say something about how that object will behave in certain circumstances, as a disposition is a power that manifests in certain circumstances. In the terminology often used in the literature, a disposition manifests when the object possessing the disposition is placed in the stimulus conditions for that disposition—the conditions which induce the disposition's causing an effect.

For example, a fragile glass has the disposition to break when struck. The glass's being struck is the stimulus condition for the disposition, and the glass's breaking is the manifestation of the disposition, which is triggered when the glass is in the stimulus condition. Similarly, a dry leaf may be flammable—having the disposition to burst into flame when it reaches a certain temperature. The manifestation is the leaf's bursting into flame, and the stimulus condition is the leaf's reaching a certain temperature.

It is clear from this that dispositions are powers that are reactive to features of their environment. An object that possesses a disposition has the power to bring about some effect, but this effect is always brought about as a response to circumstances that trigger the manifestation of the disposition. It is clear, then, why dispositions can be associated with the agency of inanimate objects. Inanimate objects bring about effects and cause changes, but we do not suppose that they do this of their own accord. Their causal activities are always due to the circumstances they are placed in.

Due to the fact that attributing a disposition to an object gives us information about how that object will behave, dispositions have often been linked to conditional statements.[1] But although there appears to be some connection between dispositions and conditional statements, attempts to analyse the notion of disposition in terms of a simple conditional statement are widely thought to be incorrect. A number of well-known counterexamples have shown that there is more to an object's having a disposition than its being the case that the object will exhibit some behaviour in certain circumstances.[2]

One classic counterexample involves a disposition being prevented from manifesting when in its stimulus conditions.[3] Imagine a glass that is fragile: it has the disposition to break when struck. On a simple conditional analysis of dispositions, what it is for the glass to have this disposition is for it to be the case that if the glass were struck, it would break. The problem for this analysis comes when we imagine wrapping the glass up in protective packaging to prevent it from breaking. The fact that the glass is wrapped up does not mean that it is no longer fragile: the packaging does not change the nature of the glass, and does remove its disposition. But now the conditional statement is no longer true. If the glass were struck it would not break, since the packaging would protect it. This and other counterexamples show that possessing a disposition is a matter of more than just behaving in a certain way in certain circumstances.

1 Proponents of the view that dispositions can be analysed in terms of conditional statements include Ryle (1949) and Quine (1960)

2 (Bird 1998); (Martin 1994); (Lewis 1997)

3 (Bird 1998)

Some philosophers use the terms 'disposition' and 'power' as if they meant the same thing.[4] This is because they take it that all powers an entity might have are powers that are manifested in response to certain conditions—in other words, all the powers there are, are dispositions. However, I am not using the term 'disposition' in this way. There are some powers that, at least on first examination, need not be manifested in any particular conditions. These are abilities, which are exercised at will rather than triggered to manifest in certain circumstances. My use of the terms 'disposition' and 'ability' is to distinguish the powers with these features: those that are manifested in certain conditions and those that are exercised at will, respectively.[5]

1.2 Abilities

Unlike the attribution of a disposition, the attribution of an ability to an agent does not by itself give any information about what an agent will do in any particular circumstances.[6] Abilities are exercised at will, rather than being manifested when the agent is exposed to stimulus conditions. An agent may have the ability to kick down a door but, even when conditions are right for the successful exercise of that ability and the agent is faced with a kickable door, the agent's ability may go unexercised.

The notion of ability plays a central part in some philosophical discussions of human agency, such as the debate around free will, responsibility and the ability to do otherwise. And, although we may not use the terminology employed in philosophical discussions of ability, it is also central to our everyday statements about human agency. We pay to see people who have the abilities to play musical instruments, and it might be that we do not understand the train conductor when we are on holiday because we are not able to speak Spanish. We might say 'She can run faster than

4 E.g. (Molnar 2003)

5 As we will see in Section 2.3 of this chapter, there have been efforts to provide an analysis of ability that reduces it to disposition. If successful, this would show that all powers are in fact manifested in response to stimulus conditions. These accounts also take as a starting point the apparently distinct natures of abilities and dispositions. So, my use of 'disposition' and 'ability' to refer to two types of powers does not presuppose the falsity of these accounts.

6 However, the attribution of an ability taken along with our awareness of other information about an agent *may* give us information about what the agent is likely to do in certain circumstances. If we know that Stuart is afraid of spiders, the information that Stuart has the ability to leap up onto tables may make us think that Stuart is more likely to do so when he is near a desk and sees a spider.

him', 'I couldn't kick a football straight before, but now I can', or 'You can carry it, it's not that heavy!' In all of these cases, the key idea in some form or another is that human agents have abilities, and that they are exercising their abilities when they act.[7]

The way I am using 'ability' here needs further clarification, since there are a number of ways in which the term is used in everyday language. Sentences like 'She can pick a lock, 'She is able to pick a lock' and 'She has the ability to pick a lock' may be understood as expressing the same thought. But each of the terms 'can', 'able' and 'ability' are context sensitive and have different connotations, leaving open the possibility of confusion about the intended target of the term 'ability'.

In some cases 'She can X' or 'She is able to X' may be synonymous with 'she has the ability to X'—indicating that the agent has a power to X that is exercised at will. However, in other cases 'She can X' may be used in a way that does not attribute an ability but simply indicates that it is possible for someone to do something.

Take a case in which Barbara, a beginner golfer, manages by some fluke, on her first try at hitting the ball, to get a hole in one. On the notion of 'ability' that I am attempting to delineate, it would not be correct to say that Barbara has the ability to get a hole in one. Possessing an ability to do something requires that there is a certain level of reliability with which one can do it. We cannot see an agent as having a power that can be exercised at will, if they do not have a sufficient level of control over whether they do what the purported power is a power to do. Barbara's hole in one is a case like this: it is simply the result of luck and is not something that she could reliably repeat. She does not have the ability to score a hole in one.

Despite the fact that Barbara lacks the ability to get a hole in one, it may nevertheless be correct to say 'Barbara can get a hole in one' or 'Barbara is able to get a hole in one'. 'She can do X' and 'She is able to do X' are sometimes used simply to mean 'it is possible for her to do X'. Just as we might say 'that cardboard box can catch fire', meaning that it is possible for the box to catch fire, we might say 'Barbara *can* get a hole in one' meaning that it is possible for her to score a hole in one. But this would

7 The notion of ability is sometimes distinguished from other powers by the fact that its exercises are actions. E.g. (Maier 2014): 'A power is an ability just in case it relates an agent to an *action*.'

not be to attribute to her the *ability* to get a hole in one. We are not attributing to her a power that can be exercised at will.

So, we must take care when looking at the language we use to indicate ability. To say that an agent possesses an ability is unambiguous in attributing a power to an agent—but this more technical way of referring to abilities is not often used in everyday speech. We more often use ambiguous phrases like 'She can climb that wall' or 'He isn't able to do that'—where we must look at the context to determine whether what is being attributed is an ability or simply the possibility of something occurring.

1.3 Which agents possess dispositions or abilities?

Before moving on, some further clarifications are needed. I have claimed that the notion of ability is central to our thought about human agency, and that the notion of disposition is central to our thought about the agency of inanimate objects. But two questions arise: do human agents have dispositions, and do inanimate objects ever have abilities?

To answer the first question: human agents do possess dispositions as well as abilities. A person may have the disposition to feel pain when pricked with a pin, or the disposition to get burnt when out in the sun too long, or to feel hungry if they have not eaten anything for a long time. In these cases, the manifestation is triggered by certain conditions, in the same way as the dispositions of inanimate objects discussed above. The manifestations of these dispositions are not actions.

There are other cases in which we may say that a human agent has a disposition, where what we mean to attribute is somewhat different to the dispositions we have been discussing.

We might say that someone has an angry disposition or a nervous disposition—thereby claiming that the agent is likely to act in certain ways in certain circumstances. Or we might say that an agent is disposed to answer the door when the doorbell rings, or to drive dangerously fast when she is late. In these cases, we are giving more information about what an agent will do in certain circumstances

than when we attribute an ability, but less information than when we attribute a disposition of the kind that may also be possessed by an inanimate object. We do not think that a person with the disposition to pick up the phone when it rings will *always* pick up the phone when it rings and no obstacles to doing so are present. Rather, we think that the agent will *usually* pick up the phone.

So, although we do attribute to human agents dispositions to act in certain ways, when we do so we are attributing characteristics somewhat different to the dispositions we normally assign to inanimate objects. My focus in the rest of this chapter will be on the type of dispositions that may also be attributed to inanimate objects, rather the features we attribute with sentences such as 'he is disposed to drum his fingers on the desk when he is getting impatient.'

To answer the second question: inanimate objects cannot possess abilities. This may sound like a strange claim, since we often say things like 'that acid can dissolve the zinc' or 'that saw can cut through metal'. But to see what is going in these cases we must remember the ambiguity of the term 'can'. With these sentences, it seems that it is a disposition being attributed to the objects, rather than an ability. I will discuss this type of example in more detail in the next section.

2.

Accounts of Disposition and Ability

2.1 Necessity

The key difference between abilities and dispositions is that abilities are exercised at will whilst dispositions have specific conditions in which they are triggered to manifest. This distinction mirrors the fact that we see the agency of inanimate objects as rigid and reactive, while we see human agents as active and flexible in what they do. Some have understood this distinction in terms of dispositions being necessitated to manifest in certain conditions, while the exercises of abilities are not

necessitated to manifest in any conditions—and have taken this to show that there is a deep divide between our concepts of disposition and ability.[8]

The difference between our concepts of disposition and ability is not so pronounced as this, however: it is not the case that our notion of disposition is the notion of a power that is necessitated to manifest in certain conditions whilst our notion of an ability is the notion of a power whose exercises are not necessitated. One way in which it could be true that this difference does not exist is if the exercises of abilities, like the manifestations of dispositions, are necessitated by certain circumstances. Some philosophers have attempted to show just this by analysing abilities in terms of dispositions. They argue that to possess an ability is simply to possess a disposition with a certain kind of stimulus conditions.

In the remainder of this chapter I will examine this approach. After introducing the views that separate dispositions and abilities on the grounds that dispositions are necessitated to manifest by certain conditions, I will examine accounts by Kadri Vihvelin and Michael Fara that attempt to avoid this separation by analysing abilities in terms of dispositions.[9] We will see that these analyses do not succeed: both depend on seeing an agent's trying to act as a stimulus condition for a disposition, a move which creates a number of problems. However, putting this approach to one side will pave the way for looking at a more fruitful approach in Chapter 3.

2.2 Separating abilities and dispositions

Richard Taylor and Peter van Inwagen both compare our disposition and ability concepts and conclude that the two notions are fundamentally different.[10] They compare sentences like 'The acid can dissolve the zinc' and 'I can move my finger', the former of which attributes a disposition to the acid, and the latter of which attributes an ability to the agent. Both philosophers take it that despite grammatical similarity of sentences such as these, there is no prospect for assimilating abilities to a more general category of disposition—and that this is because the manifestations

8 (Taylor 1968) Ch. 4; (van Inwagen 1983)

9 (Vihvelin 2013); (Fara 2008)

10 (Taylor 1968) Ch. 4; (van Inwagen 1983)

of dispositions are necessitated by certain circumstances whilst the exercises of abilities are not.

The two philosophers have different views of the nature of disposition, but both take the idea that manifestations are necessitated by certain conditions to be a core feature of our concept of disposition.[11]

Taylor's account of disposition is similar to the simple conditional analysis discussed in Section 1.1. He argues that a sentence like 'the cricket ball can smash a window' expresses the idea of a causal connection between certain events such that if one were to occur, then, necessarily, the other would follow.[12] The sentence expresses a relationship between an event of the cricket ball hitting the window and an event of the window smashing such that if the former event were to occur, the latter would necessarily follow.

Van Inwagen remains largely silent on his understanding of the 'disposition' concept. He is dubious about the merits of conceptual analysis, and states that he is content to rely on some intuitive understanding of 'disposition'. But it is still clear that, like Taylor, van Inwagen takes the idea of a manifestation being necessitated as central to the notion of disposition. He writes: 'the concept of causal power or capacity is the concept of an invariable disposition to react to certain determinate changes in the environment in certain determinate ways.'[13]

I will look at two arguments proposed by Taylor and one by van Inwagen for the separation of dispositions and abilities. As we have seen, Taylor understands sentences like 'the acid can dissolve the zinc' as expressing a connection between two possible events, such that if one were to occur the other must follow. In this case we can specify what those two events might be: the zinc's coming into contact with the acid and the zinc's dissolving. Taylor argues that if our notion of ability employed a similar idea we should expect to find two events that could feature in an analogous paraphrase of a sentence like 'I can move my finger'. But we struggle to find two such events.

---

11 Neither Taylor nor van Inwagen use the term 'disposition'—both use the terms 'power' and 'capacity' to refer to their target. I take this to be a difference in terminology only: the powers that they discuss are the same as those we have termed dispositions.

12 (Taylor 1968) p. 45

13 (Van Inwagen 1983) p. 11

Taylor considers the possibility of paraphrasing 'I can move my finger' as a statement about the relation between a possible internal physical event and the event of my finger moving: for example, 'if a certain muscle were to contract my finger would move' or 'if a certain event in my brain were to occur my finger would move.' But he concludes that these sentences do not communicate the idea of an agent exercising a power that is communicated by 'I can move my finger'. An event in my brain or a muscle movement may cause my finger to move even if I am not exercising an ability to move my finger—my finger's movement may simply be a muscle twitch or spasm.

Another possibility for paraphrasing 'I can move my finger' is considered by Taylor. He asks whether we might paraphrase the sentence as a statement about the relation between a possible mental event and the event of my finger moving: 'I will move my finger if I want to' or 'I will move my finger if I choose to do so'. This option is problematic as well—Taylor argues that so-called events like wanting or choosing 'have the semblance of fiction.'[14] There is no reason to posit any mental event corresponding to wanting, choosing or intending unless one is already trying to fit these notions into an event-causal framework.

A second argument by Taylor is intended to show that there is a disanalogy between the kinds of cases in which sentences attributing dispositions and those attributing abilities are true. Taylor asks us to imagine a particular volume of acid that has the disposition to dissolve zinc. The acid is locked away in a safe so that it cannot come into contact with zinc. Still, it seems that 'the acid can dissolve zinc' remains true. Taylor sets out an analogous situation in the case of human action: my hand is placed in a cast so that my fingers are rendered immobile. But in this case, Taylor argues, it seems that 'I can move my finger' is false.

Taylor claims that the comparison between these two cases shows that the notions of ability and disposition are very different, and must be analysed in different ways. We can explain why it is true to say 'the acid can dissolve the zinc' even when the acid is locked in a safe by noting that the sentence is equivalent to '*if* there were an event of the zinc coming into contact with acid, there would follow an event of the acid dissolving the zinc'. But the fact that it is false to say 'I can move my finger'

14 (Taylor 1968) p. 50

when my hand is in a cast shows that this sentence attributes something other than a link between two possible events.

The third argument is made by van Inwagen. He focuses straightforwardly on the idea that dispositions are necessitated to manifest by certain conditions, whilst the exercises of abilities are not necessitated to manifest:

> *For a man to have the [disposition] to understand French is for him to be such that if he were placed in certain circumstances, which wouldn't be very hard to delimit, and if he were to hear French spoken, then, willy-nilly, he would understand what was being said. But if a man can speak French, then it certainly does not follow that there are any circumstances in which he would, willy-nilly, speak French. The concept of a causal power [...] would seem to be the concept of an invariable disposition to react to certain determinate changes in the environment in certain determinate ways, whereas the concept of an agent's power to act would seem not to be the concept of a power that is dispositional or reactive, but rather the concept of a power to originate changes in the environment.*[15]

The idea here is that when a sentence like 'the acid can dissolve the zinc' is true, this means that if the right circumstances held and all obstacles were removed, the acid would dissolve the zinc. But when a statement such as 'I can move my finger', is true, it is not the case that I must move my finger if I am in the right circumstances and all obstacles are removed. When I have the ability to move my finger, it is up to me whether I move it or not.

2.3 Linking abilities and dispositions

Kadri Vihvelin and Michael Fara propose very similar analyses of abilities, which purport to show that abilities are necessitated to manifest in certain conditions in much the same way as dispositions.[16] Both philosophers' main focus is free will. They aim to show that human agents' possessing 'the ability to do otherwise' is compatible with determinism by demonstrating that our notion of ability is adequately captured by an account on which the exercises of abilities are necessitated

15 (van Inwagen 1983) p. 12. van Inwagen originally uses the terms 'causal power' and 'capacity'. I have changed these to 'disposition' in the quotation so as not to confuse terminology, as it seems that despite the difference in terminology, van Inwagen is discussing the power that we are calling disposition.

16 (Vihvelin 2013); (Fara 2008)

by certain conditions. However, both philosophers take it that they are providing analyses of the notion of ability in general—so their focus on free will does not preclude their accounts' importance for the purposes of this chapter.

Taylor's, and to a lesser extent van Inwagen's, arguments that ability and disposition are fundamentally separate draw on a conditional understanding of disposition. They focus on the idea that, while sentences attributing dispositions can be paraphrased as a statement about what an object will do in certain circumstances, sentences attributing abilities cannot fit into this form. Both Vihvelin and Fara reject the simple conditional analysis of disposition in favour of more complex accounts. They believe that these more complex analyses give them the tools to show that ability can be analysed in terms of disposition.

Before looking at the two accounts, it is important to clarify exactly what Vihvelin's and Fara's target is. There are two senses of ability that a sentence like 'I can move my finger' could be referring to. I will term these senses 'general' and 'specific'.[17] Taylor and van Inwagen fail to make the distinction between these different senses, but Vihvelin and Fara are aware of the distinction and make it clear that their target is general ability.[18]

We can begin to understand the distinction between these two senses of ability by looking at an example. Steven has the ability to ride a bike—and his possession of this ability is not affected by whether he can exercise it *at this very moment*. It just so happens that Steven is climbing Mount Everest, so there are no bikes in the immediate area and there is no possibility that he could right now exercise his ability to ride a bike—but it still seems right to say that he has this ability. The sentence 'Steven can ride a bike' could express this. However, in exactly the same situation it would also be true to say 'Steven cannot ride a bike'—meaning that he cannot *right now* ride a bike, because he lacks the opportunity to do so.

17 Vihvelin (2013) draws the same distinction using different terms. She uses the term 'narrow ability' where I use 'general ability', and the term 'wide ability' where I use 'specific ability'. I will use the terms 'general' and 'specific', as these are more widely used in the literature.

18 (Vihvelin 2013, Ch. 6); (Fara 2008, pp. 844-845). Unlike Vihvelin, Fara does not explicitly draw this distinction. However, he states that he is interested in the abilities that an agent may continue to possess even when the agent lacks the opportunity-—i.e. general abilities.

In the former sense, 'Steven can ride a bike' is being used to refer to a general ability to ride a bike. An agent's possessing a general ability is a matter of the agent having certain intrinsic features—such as skills, competencies and physical characteristics. An agent's possessing a general ability at a time does not depend on the agent having the opportunity to exercise the ability at that time. In the latter sense, 'Steven can ride a bike' is referring to a specific ability to ride a bike. An agent's having a specific ability is a matter of their having a general ability *and* having the opportunity to exercise that ability. Steven possesses the general ability but lacks the specific ability in the situation described because he lacks the opportunity to exercise his ability. But in a situation where there is a bike right in front of him he would possesses the general ability *and* the specific ability.

Vihvelin and Fara make clear that their target is general ability—and it is right that general ability should be our focus. Although we talk about both specific and general abilities in everyday life, the notion of general ability is conceptually prior to the notion of specific ability—the latter notion rests on the former. Our notion of general ability is simply the notion of a power exercised at will that we have been referring to as 'ability' in the above discussions. To attribute a specific ability to an agent is simply to attribute a general ability that the agent has the opportunity to exercise at that time. Once we have an account of general abilities, giving an account of specific abilities is a simple matter.

We can now look at Vihvelin and Fara's accounts of ability. Despite their similarities, the two accounts link abilities to dispositions in different ways. While Fara argues that having an ability entails having a disposition, Vihvelin argues that abilities *are* dispositions or collections of dispositions.

Fara argues that an agent has an ability to A in circumstances C if and only if the agent is disposed to A when, in C, she tries to A.[19] Both the ability and the disposition in this formula have wide scope reading: Fara intends his analysis to refer to the ability to [A in circumstances C] and the disposition to [A, when, in C, the bearer of the disposition tries to A]. For Fara, possessing an ability requires possessing a disposition with a certain sort of stimulus condition—part of the stimulus condition must be the agent's trying to A.

19 (Fara 2008) p. 848

Vihvelin's dispositional analysis of ability is roughly based on David Lewis' analysis of dispositions. Lewis attempts to improve on the simple conditional account and avoid counterexamples like Bird's by suggesting that possessing a disposition requires an object to have certain intrinsic properties.[20] Vihvelin draws on this picture of dispositions, and, like Fara, adds the notion of the agent's trying to do something:

> *S has the ability at time t to do X as the result of trying iff, for some intrinsic property B that S has at t, for some time t' after t, if S had the opportunity at t to do X and S tried to do X while retaining property B until time t', then, in a suitable proportion of these cases, S's trying to do X and S's having of B would jointly be an S-complete cause of S's doing X.*[21]

Both Vihvelin and Fara have accounts of disposition which move away from the simple conditional analysis, and which they take to avoid the standard counterexamples that often afflict analyses of disposition. It will not be necessary to look in detail at their accounts of disposition in order to decide whether their dispositional analyses of ability succeed. It is not on the detail of their accounts of disposition that their analyses of ability will stand or fall, but on their inclusion of the agent's trying as a stimulus condition for the agent's powers. It is through this notion of trying that they hope to account for the fact that abilities are exercised at will, but it is far from clear that the agent's trying can play the role that it is intended to on Fara's and Vihvelin's analyses.

We will look closely at the notion of trying in Section 3 of this chapter. First, we will look at one of the arguments put forward by Taylor which be dealt with straightforwardly by Fara and Vihvelin. While the notion of trying is required to overcome the other two arguments set out in Section 2.2, it is not required in this case. All that is required to show why this argument of Taylor's fails is a proper understanding of the distinction between general and specific abilities.

20 (Lewis 1997)

21 (Vihvelin 2013) p. 188

2.4 Replying to Taylor's argument

The distinction between general and specific abilities can allow us to deal with Taylor's argument that there is a disanalogy between the situations in which statements ascribing disposition and statements ascribing ability are true. Recall that Taylor uses the example of a volume of acid kept in a safe, and compares it to an analogous case where a person's fingers are restricted by a cast. He argues that it seems right to say that the acid retains its disposition to dissolve zinc even when the acid is locked away in a safe, but it seems false to say that I still have the ability to move my finger when it is restricted by a cast. Taylor intends this to show that in stating that the acid has a disposition to melt zinc, we must be ascribing a different kind of power to that which we ascribe in stating that the agent has an ability to move her hand.

Employing the distinction between general and specific abilities makes it clear that the two cases are in fact analogous. Taylor claims that I lack the ability to move my finger if it is restricted by a cast—but, with the distinction in mind, we can now see that what I lack is the *specific* ability to move my finger. When I have a cast on, I no longer have the opportunity to move my finger, so it is false to say 'I can move my finger' if we mean this to attribute a specific ability. I have not lost my general ability, however—this ability is due to my intrinsic properties and my possession of this ability does not depend on whether I right now have the opportunity to exercise it. Just as the acid retains its disposition to dissolve the zinc even when it is locked away in a safe, I retain my general ability to move my finger even when my finger is restricted by a cast.

But why does it seem natural to think that 'the acid can dissolve the zinc' is true when the zinc is in the safe, while it seems natural to think that 'I can move my finger' is false when my finger is restricted by a cast? This seems to be because the most natural understanding of the sentence 'I can move my finger' is as referring to the possession of a specific ability to move my finger, rather than a general ability. Perhaps because we take it for granted that the general ability to move one's finger is possessed by most people, we take the sentence 'I can move my finger' to be communicating something more informative by referring to the specific ability. Nevertheless, when we clarify what is meant by this statement, by employing the

distinction between general and specific abilities, we can see that it is correct to attribute a general ability in this case. Taylor has not shown a disanalogy between our notions of disposition and ability.

3.

Trying

3.1 Replying to further arguments by Taylor and van Inwagen

Whether or not Vihvelin and Fara can answer the other challenges posed, and successfully demonstrate that ability can be analysed in terms of disposition, turns on the inclusion of 'trying' in their accounts.

As we have seen, what leads van Inwagen to claim that abilities and dispositions are distinct kinds of powers is that dispositions are necessitated to manifest by certain circumstances whilst the possessors of abilities are not necessitated to exercise their abilities in any circumstances. His thought is that it is extremely implausible to think that an agent's having the ability to do something like throw a frisbee, or to kick a football involves there being some situations in which the agent is *must* throw a frisbee or kick a football. An agent's possessing an ability to do something, unlike an agent's possessing a disposition, does not involve there being circumstances in which the agent must do that thing.

It is by including 'trying' in their accounts that Vihvelin and Fara hope to avoid this problem. They aim to show that the exercises of abilities are in fact necessitated by certain circumstances. But, whilst van Inwagen is thinking only of circumstances external to the agent, Vihvelin and Fara propose that one of the circumstances that necessitates an agent's exercising her ability is the 'internal' circumstance of the agent's trying to act. So, contra van Inwagen, understanding abilities in the same way as dispositions does not entail that someone with the ability to speak French must speak French 'willy-nilly' in certain circumstances. A person with the ability to speak

French will speak French only if she tries to speak French (and other favourable conditions are present).

This appears to avoid the problem raised by van Inwagen, but it is not clear that Vihvelin and Fara's proposal that trying is a stimulus condition for a power is itself a plausible one. Recall Taylor's first argument, in which he attempts to find two events that could fit into a conditional statement to paraphrase 'I can move my finger'. Taylor considers paraphrasing the sentence as 'I will move my finger if I want to' or 'I will move my finger if I choose to do so', but decides not to pursue this path due to doubts about the existence of supposed mental events like wantings and choosings. Trying falls into the same category, and it is likely that 'I will move my finger if I try to so' would also be rejected by Taylor on the same grounds.

To determine whether Vihvelin's and Fara's accounts of ability are successful, we must look more closely at what they take 'trying' to be, and at whether it is something that could play the role of a stimulus condition for a disposition. We will see that on any plausible view of what trying is, it is difficult to see how it could play this role—and that there are further problems caused for Vihvelin and Fara by the inclusion of trying in their analyses of ability.

3.2 What is trying?

Let us look in turn at how Vihvelin and Fara understand an agent's trying to do something. We will start with Vihvelin's view of trying, which is somewhat difficult to uncover. It is clear that Vihvelin aims to distance herself from the view of trying as some prior mental event—the view of the nature of trying that would be rejected by Taylor. But it is less clear what her positive conception of trying amounts to.

Vihvelin writes:

> *We try to do X whenever we acquire an intention or desire to do X, here and now, and that intention or desire causes at least the beginning of the process of*

> *doing something that we believe, perhaps mistakenly, will move us closer towards our goal of doing X.*[22]

There are two ways of reading Vihvelin's remarks about trying. On one reading her picture of the nature of trying seems implausible, but on the other reading her remarks are uninformative about how we should understand the nature of trying.

On one reading we can take Vihvelin's remarks to be a claim about what trying *is*: trying is just beginning to act—starting the process of acting—and if one has the ability to act in that way, one will usually be successful when one tries. But Vihvelin cannot be claiming that trying *is* beginning to act because this conflicts with her other claim that trying causes actions. She believes that trying is an event or sequence of event that causes 'the bodily movement events that some philosophers say are our actions'.[23] An agent's starting to act cannot cause the event of his acting, because the event of his acting will already have begun if he is starting to act. Starting to move my finger cannot cause the event of my finger moving—because if I am starting to move my finger, my finger is already moving.

This leaves us with the second, uninformative reading. On this reading, we can understand Vihvelin to be simply making a claim about the situations in which trying normally occurs: we try to act when we have an intention to act and that intention causes the beginning of the process of our acting. But if we think, as is widely accepted, that an agent always tries to do what they do intentionally, and then it is obvious that trying occurs in the situation described by Vihvelin. This does not help us get any closer to an understanding of what trying is and how it could play the role of a stimulus condition for a disposition.

Fara's view on the nature of trying is slightly clearer. He, too, avoids the view that trying is a prior mental event. Trying to act, for Fara, is 'placing oneself, as best one can, in circumstances appropriate to or conducive to performing the action'.[24] To illustrate this, he gives the example of a person trying to get to sleep by doing long division in her head. His thought is that this person's trying amounts to her doing something dull to create conditions conducive to her falling asleep.

---

22 (Vihvelin 2013) pp. 176-177

23 Ibid. p. 177

24 (Fara 2008) p. 849

There are two problems that arise from Fara's and Vihvelin's inclusion of trying in their accounts of ability.

3.3 The first problem: a regress of tryings

On Vihvelin's and Fara's accounts, the possession of an ability requires the possession of a disposition that has trying to act as a stimulus condition. The first problem presents itself when we ask whether trying to do X requires that the agent possesses the *ability* to try to do X. If trying does require the ability to try, then according to Fara's and Vihvelin's analyses of ability, this ability to try requires a disposition to try that has trying to try as a stimulus condition. The same question can then be asked of trying to try: does trying to try require the ability to try to try? It seems that Vihvelin and Fara are in danger of introducing a regress.

Whether there really is a problem for Vihvelin and Fara depends on two things: whether the regress actually follows from their accounts of ability, and to what extent a regress of tryings is unacceptable. In considering the problem I will focus mainly on Vihvelin's view. This is because, whilst Fara does not discuss this potential problem for his account, Vihvelin does attempt to address the issue.

Vihvelin's first line of defence against the problem is to claim that the regress does not actually follow from her view. She argues that she is not committed to the claim that trying to act requires the ability to try. Vihvelin claims that within her view that the ability to act has trying as a stimulus condition there is nothing that entails that the agent must also possess the ability to try:

> *Having a disposition to give response R to stimulus S doesn't entail having any additional disposition to give S as a response to some prior stimulus. I say that to have the narrow ability to do S is to have the disposition to do X in response to the stimulus of trying to do X; therefore, I say that having the narrow ability to do X doesn't entail having any additional ability (disposition) to try to do X in response to some prior stimulus.*[25]

Vihvelin is right in some ways: her claim that the ability to act has trying as a stimulus condition does not by itself entail that the agent must have the ability to try

25 (Vihvelin 2013) p. 176

to act. There is no general rule concerning dispositions that will get us from the former claim to the latter. But the idea that trying to X requires the ability to try to X does not come from our belief in such a rule. There are other reasons to think that Vihvelin and Fara must accept that an agent's trying to X must be the exercise of an ability to try to X.

The role of trying in Fara's and Vihvelin's analyses of ability is to account for the idea that whether or not we exercise our abilities is up to us. With the idea that trying is a stimulus condition, we can accept that the exercises of abilities are necessitated in certain circumstances without being led to the conclusion that agents find themselves exercising their abilities 'willy-nilly' in a given situation. But key to this is the thought that trying is itself something that is done at will. It is not clear that trying would be able to play the role of introducing an element of voluntariness to the conditions for the exercise of ability if we did not already believe that trying to do something itself requires an ability that is exercised at will. So, Vihvelin and Fara appear to be committed to the claim that trying to act requires the ability to try.

Vihvelin would dispute this line of thought—she argues that trying *can* account for the idea that abilities are exercised at will even if trying to act is not an exercise of the ability to try. She argues that it is due to the kinds of abilities that we have that it is up to us what we do, rather than the way the exercise of these abilities is triggered.

Vihvelin gives the example of a very young child crawling across the room to grab her toys.[26] In a case like this, the child is acting and her bodily movements are made in response to her trying to do something, which is in turn a response to her perceptual beliefs about the objects in front of her. Vihvelin argues that it would be a mistake to think that the child is not in control of her actions simply because her trying and thereby her exercising of her abilities is triggered by beliefs and other mental states.

Vihvelin's point seems to be that it is not clear what more we could want from an account of ability than what her account gives us. She is suggesting that those who argue that an ability is not exercised at will unless the trying that is its stimulus condition is also the exercise of an ability are themselves being led towards an

26 (Vihvelin 2013) pp. 178-179

implausible view. If they are arguing that abilities are not exercised at will because their exercise is the result of certain conditions, like the presence of trying or belief and desire, then do they instead want an account on which the exercise of ability is not due to any conditions at all, and is completely spontaneous? Vihvelin is arguing that we clearly must accept that exercises of abilities are necessitated by certain conditions: what makes the difference to whether or not we find ourselves with a picture on which agents do things 'willy-nilly' is whether the conditions that we take to necessitate these exercises are external to the agent or are 'internal', like the agent's trying.

This response from Vihvelin is too weak to overcome the problem. Vilhvelin attempts to defend her account by appealing to the absurdity of the only alternative position. But we are only restricted to a choice between her view and the absurd alternative if we are already constrained to conceiving of ability as a disposition with trying as a stimulus condition. There is another option in addition to Vihvelin's view and the absurd alternative: rejecting the conception of ability as a disposition in the first place. Vihvelin's response leaves it open that we should simply reject her view entirely, and deny that abilities can be analysed in terms of dispositions.

We have seen, then, that there is no satisfactory way for Vihvelin and Fara to deny that trying to act itself requires the ability to try to act. They are committed to a regress of tryings—for trying to try to act will also require the ability to try to try to act in order to preserve the voluntariness of ability, and trying to try to try to act will require the ability to try to try to try to act, and so on. As mentioned above, another way out of this problem for Vihvelin and Fara is to accept that the regress occurs, but deny that it is vicious. Perhaps the existence a regress of tryings, each the exercise of an ability, is something we should simply accept.

Vihvelin also proposes this defence against the problem. She claims: 'it makes perfectly good sense to say that someone is trying to try to do something.'[27] And this claim does appear to be correct. Imagine a case where I want to demonstrate to a friend that a door is locked. My aim is to demonstrate this by trying to open the door—when the door does not open my friend will see that it is locked. When I perform my action of trying to open the door (and assuming that whenever a person

27 Vihvelin 2013) p. 179

does something intentionally they are trying to do it) I am trying to try to open the door.

It may be the case that trying to try makes sense, but Vihvelin makes no mention of trying to try to try, and all of the other tryings that we have seen are entailed by her account. It still seems implausible that all of these tryings are present whenever someone acts, and for these reason the regress still seems unacceptable.[28]

The most serious problem indicated by the regress is that the analyses proposed by Vihvelin and Fara are circular. They aim to show how we can see abilities as being exercised at will, whilst at the same time accepting that the exercises of abilities are necessitated by certain conditions. But Vihvelin and Fara attempt to achieve this by introducing as one of the conditions that necessitates the exercise of an ability something which itself requires the exercise of ability. They cannot claim to have successfully shown that the notion of ability can be accommodated by an analysis in terms of dispositions—their accounts depend on their introducing the notion of ability back into the analysis.

3.4 The second problem: trying as a stimulus condition

The second problem with the use of trying in the analyses of ability is that, on a correct conception of the nature of trying, it does not appear that the agent's trying to do something can play the role of a stimulus condition for a disposition. If trying is understood as a mental event that occurs prior to the action, then it is relatively straightforward to see how it could be a trigger for the exercise of an ability. But, as we have seen, both Vihvelin and Fara rightly attempt to distance themselves from this problematic conception of trying.

I will focus here on Fara's understanding of trying. As we saw above, Vihvelin does not offer us a clear picture of the nature of trying that goes deeper than our common-sense understanding of the concept. Fara gives a slightly more detailed account. On his view, one tries to do something by putting oneself, as best one can,

28 In setting out his own account, Fara claims that 'one cannot, even in principle, try to try to sing a song.' (2008, p. 849)

in circumstances conducive to performing the action. To illustrate his notion of trying, he uses the example of someone trying to get to sleep by doing long division in her head. Fara's thought is that by doing long division in her head, the person is putting herself in circumstances conducive to falling asleep by making herself relaxed.

The problem with this conception of trying is that putting oneself in 'circumstances conducive to performing the action' seems a very indirect route to action. On this view, when we are trying to do something we are simply placing ourselves in certain circumstances and hoping that the action will follow. We can see how implausible this is by noting that it is very difficult to see how this idea of putting oneself in circumstances conducive to action could even apply to cases where an agent is acting. When I try to do something like opening a door, I am not trying by putting myself in circumstances conducive to opening a door unless by this we just mean that I am doing what it takes to open the door: turning the handle and so on. The example Fara uses to illustrate his view—trying to fall asleep—does not concern an action at all: falling asleep is something that happens to the agent. The fact that Fara does not give an example of his view of trying that concerns an action perhaps illustrates its unsuitability.

Fara's understanding trying in this way may be influenced by his need to show how trying can be a stimulus condition for a disposition. An agent's being placed in circumstances conducive to performing an action does seem appropriate to being a stimulus condition for the manifestation of a disposition. But, while it may serve the purposes of the analysis of abilities, Fara's view does not reflect the way we conceive of our trying to do something.

Jennifer Hornsby argues for a conception of trying that sounds very similar to Fara's: Hornsby proposes that 'to *try* to Φ is to *do what one can* to Φ.'[29] Perhaps this is what Fara has in mind for the way in which his conception of trying could apply to cases of action. On this view, my trying to open the door *just is* my turning the handle, pushing it with my shoulder if it is stuck, and so on. This appears much more direct and therefore more plausible than Fara's picture of trying as taking the indirect route of 'putting oneself in circumstances conducive to action'.

29 (Hornsby 2010)

But, although Hornsby's view offers a way to interpret Fara's view of action as more plausible, Fara cannot accept it. If he were to accept Hornsby's view, he would again be faced with the problem that trying cannot play the role of a stimulus condition for the exercise of an ability. When I am trying to open the door, by turning the handle and so on, my action of opening the door has already begun—I am already in the process of opening the door. Trying to open the door cannot trigger the exercise of my ability on Hornsby's conception of trying, because if I am trying to open the door my action has already begun.

3.5 The problem with trying

Although the inclusion of trying in the analysis of ability is Vihvelin's and Fara's main tool in their aim to show that abilities can be reduced to dispositions, we can now see that this move should be rejected due to the problems that the introduction of trying creates. We have seen that the two main problems Vihvelin and Fara face are that the agent's trying to do something cannot be a stimulus condition for the manifestation of a disposition, and that appealing to the notion of trying in their accounts leads to circularity.

In the case of the former problem, they have a choice between presenting a picture of the agent's trying to do something on which it can be a stimulus condition for a disposition—which leads to a view of trying implausibly far removed from our everyday understanding of it; and attempting to give a picture of trying that reflects our everyday understanding, which leaves it unclear how an agent's trying to do something can be the stimulus condition for a disposition. In the case of the latter problem, they have the choice between introducing a notion of trying on which trying is the exercise of an ability—which introduces circularity into their analyses of ability; or they can attempt to introduce a notion of trying which does not presuppose that trying is the exercise of ability—which avoids circularity but then gives us no reason to think that their analysis captures the idea that agents exercise their abilities at will.

Conclusion

What may have at first seemed a promising way of diminishing the difference between agents' manifesting dispositions and exercising abilities is ultimately a dead end. The only way our concept of ability can be given a dispositional analysis, is if the analysis employs a notion of trying that presupposes the very notion of ability that the analysis is aimed at. We should therefore abandon this line of inquiry, and look elsewhere for the way to show that the difference between disposition and ability does not lie in necessity.

Chapter 3

Disposition and Necessity

In Chapter 2 we looked at attempts to show that the exercises of abilities are necessitated by certain conditions, and concluded that these attempts fail. In this chapter we will look at a more successful way of showing that the difference between ability and disposition does not lie in necessity. We should not attempt to demonstrate that exercises of abilities are necessitated by certain conditions. Rather, the reason that the difference between ability and disposition does not lie in necessity is that neither the exercises of abilities nor the manifestations of dispositions are necessitated in any circumstances.

Conditional analyses of disposition are to some extent responsible for the idea that these powers' manifestations are necessitated by certain conditions: on these accounts, what it is for an object to have a disposition to do something is its being the case that if some condition held, the object would behave in a certain way. This gives us a picture on which certain circumstances necessitate objects' behaving in certain ways. However, there are some who argue that our notion of disposition cannot be analysed. This affords them the freedom to reconsider the relation between dispositions' manifestations and their stimulus conditions.[1]

I will focus on an account presented by Stephen Mumford and Rani Lill Anjum. They argue that our notion of disposition does not involve necessity, and that dispositions are never necessitated to manifest by any conditions. There are some respects in which we must deviate from their account, but through examining their view we will see that neither dispositions nor abilities involve necessity.

Mumford and Anjum's account highlights a further feature of disposition that seems to introduce a new, fundamental difference between disposition and ability: Mumford and Anjum argue that dispositions tend toward their manifestations. This feature appears to have no place in the case of ability: our notion of ability does not

[1] For example, Erasmus Mayr (2011, pp. 161-164) argues that our notion of disposition cannot be analysed. He gives a taxonomy of different types of disposition, some of which must be manifested in certain circumstances, and other which need not manifest when their stimulus conditions are present.

involve the idea of these powers tending toward their exercise. However, I will argue that we can develop a view of this 'tending' on which it is a feature of both dispositions and abilities.

Before we begin to look at Mumford and Anjum's arguments a note on terminology is required. Mumford and Anjum take it that in giving an account of dispositions, they are giving an account of causal powers in general: they use 'dispositions' simply to refer to causal powers. I will continue to use 'disposition' and 'ability' in the ways marked out in the previous chapter. This will be useful for further discussion of these powers' features in later chapters.

1.

Disposition and Necessity

1.1 Mumford and Anjum's account of dispositions

Mumford and Anjum argue that conditional analyses of disposition fail due to the well-known counterexamples that draw on the possibility of interference with or prevention of the manifestation.[2] 'Mask', 'mimicker' and 'fink' cases, along with other varieties of counterexample, show that our disposition ascriptions cannot be reduced to conditional statements about what would happen in certain circumstances.[3] Mumford and Anjum argue that not only does the simple conditional analysis fail, but so do modifications of the analysis that introduce new requirements such as the agent's having certain intrinsic characteristics.[4]

The debate about reductive accounts of dispositions and their counterexamples follows much the same pattern as the debate around the causal theory of action and the problem of deviant causal chains, which we looked at in Chapter 1. Each new

2 (Mumford & Anjum 2011b) pp. 382-385

3 Some of these counterexamples are discussed in (Bird 1998); (Martin 1994) and (Lewis 1997).

4 Such an analysis is presented by Lewis (1997).

counterexample leads to a new, more complex analysis of dispositions which is intended to overcome this counterexample. In turn, a new counterexample is produced which defeats this new analysis, and the search begins for a new analysis which will avoid the latest counterexample. Rather than relaying this debate, I will focus on the positive account of disposition presented by Mumford and Anjum.

Mumford and Anjum argue that dispositionality is a basic modality, different from necessity or possibility. On their view, dispositions tend towards their manifestations but never necessitate them. A disposition 'is a real tendency towards one kind of outcome from the many that are logically possible'.[5] But a disposition 'never necessitates [an outcome], even in the cases where it does succeed in producing that outcome.'[6] When we ascribe a disposition, we ascribe a tendency towards certain outcomes rather than a necessary connection between certain conditions and the disposition's manifesting, or the mere possibility of certain outcomes. To see how Mumford and Anjum arrive at this conclusion, we must look in turn at their arguments distinguishing necessity and possibility from disposition.

The argument for separating dispositionality and necessity is based on the possibility of interference. Mumford and Anjum take it that if A necessitates Y, then it follows that Y must hold whenever A holds.[7] So, if stimulus conditions for a disposition necessitate its manifestation, then the disposition must manifest whenever these stimulus conditions are present. If there are cases where the stimulus conditions for a disposition are present but the disposition does not manifest, then this shows that the stimulus conditions do not necessitate the manifestation. Mumford and Anjum argue that some of the cases of interference and prevention given as counterexamples to the conditional analysis of dispositions show exactly this—they are cases where the stimulus conditions for a disposition are present but the disposition does not manifest.

If a fragile glass is covered in bubble wrap, it is no longer the case that it would break if dropped. The stimulus condition (the glass being dropped) is present, but because another condition (the glass being wrapped in bubble wrap) is present, the

5 (Mumford & Anjum 2011b) p. 381

6 Ibid.

7 Ibid. p. 386

manifestation does not occur. It cannot be the case, then, that being dropped necessitates the glass breaking: there are possible situations where the condition is present but the glass does not break. Since this possibility of interference or prevention due to the addition of another condition is always present for all dispositions, Mumford and Anjum argue that stimulus conditions of dispositions never necessitate the manifestation of the disposition. No matter the situation, there could always be added another condition that would prevent the disposition from manifesting despite its stimulus conditions being present.

Mumford and Anjum argue that our notion of dispositionality does not involve necessity, but they also aim to show that dispositionality is stronger than mere possibility. It is not merely possible that a disposition will manifest when it is in its stimulus conditions. Rather, the disposition tends towards a particular outcome from many possible outcomes. Mumford and Anjum acknowledge that dispositionality is linked to possibility in some way —if S has a disposition to A then this entails that S's bringing about A is possible—but they aim to show that it is stronger than possibility alone.

To show that dispositionality is stronger than possibility, Mumford and Anjum look at cases where it seems right to say that some change is possible for an object, but it does not seem right to say that the object has a disposition to change in that way. They argue that it is in some sense possible for a match, when struck, to turn into a chicken—but even though this is possible, a match does not have the disposition to turn into a chicken.[8]

The example of a match turning into a chicken can produce varying intuitions. It is not clear that a match turning into a chicken *is* possible, and it might seem that if it were possible it *would* be right to say that the match had a disposition to turn into a chicken in some circumstances. It is important to clarify here, however, that Mumford and Anjum are concerned to show dispositionality is stronger than *logical* possibility rather than *natural* possibility.[9] The sense in which it is possible for a match to turn into a chicken is logical possibility—a match turning into a chicken is not logically inconsistent. But in the sense of natural possibility—what is possible

8 (Mumford & Anjum 2011a) p. 181

9 (Mumford & Anjum 2011b) p. 389

within the laws of nature—it is not possible for a match to turn into a chicken. Mumford and Anjum accept that dispositionality might be 'what determines what it is for something to be naturally possible.'[10] But their claim is that dispositionality is stronger than pure logical possibility.

So, dispositionality is weaker than necessity and stronger than logical possibility on Mumford and Anjum's view. They claim that dispositionality is a basic concept and so cannot be fully explicated in non-dispositional terms. Mumford and Anjum argue that we gain the concept of dispositionality from our experience of our own agency. As we move through the world, we experience the dispositions of other things pushing towards certain outcomes and our interference with the manifestation of those powers. When I walk down the street in a strong gale, I experience the force of the wind that would knock me over if I were not struggling against it. I can feel, when I stop struggling against it, the wind's force beginning to knock me over. It is from this kind of experience that we gain the concept of dispositionality—the notion of powers pushing towards certain manifestations but not necessitating these manifestations, since there is always the possibility of other powers counteracting them.

1.2 Assessing Mumford and Anjum's argument

It is not clear that Mumford and Anjum's arguments are wholly successful in showing that dispositions are never necessitated to manifest. We can establish that our notion of disposition does not involve necessity, but this will require looking at their argument in a new way.

As we have seen, Mumford and Anjum's argument is based on the idea that if A necessitates B, then B must be present whenever A is present. So, if it is true that stimulus conditions necessitate a disposition's manifesting, it must be the case that the disposition manifests whenever these stimulus conditions are present. Mumford and Anjum argue that there is always the possibility of some other condition being present alongside the stimulus conditions, which would prevent the manifestation

10 (Mumford & Anjum 2011b) p. 389

from occurring. From this they conclude that dispositions are never necessitated to manifest.

This argument, however, does not appear to establish as much as Mumford and Anjum require for their conclusion. Mumford and Anjum's argument clearly does establish that no one condition alone can necessitate the manifestation of a disposition: the fragile glass being struck cannot necessitate that the glass breaks, since the glass might be wrapped up and fail to break even when it is struck—and there are many other things that could prevent the glass from breaking. Whenever we look at a single condition, there is always something that could be added to prevent the disposition from manifesting. When we look at groups of conditions, however, it seems that they can in fact pass the test set out by Mumford and Anjum.

It may initially seem that a collection of stimulus conditions A, B, C and D cannot necessitate a disposition's manifestation, since this collection of conditions fails the test in the same way that a single condition does. No matter how large the collection of conditions, there is always the possibility of a further condition being added that prevents the disposition manifesting. If A, B, C, D *and E* are present then the manifestation will not occur even though A, B, C and D are present—so it cannot be the case that the collection of A, B, C and D necessitates the manifestation. Salt's manifesting its disposition to dissolve cannot be necessitated by its being put into water, the water being the right temperature, there being the right amount of water and so on—because there is always something else that could prevent the salt from dissolving despite all of these conditions being present, such as the water's already being saturated with some other soluble substance.

But this is not the only way to think about groups of conditions. On a different view of the individuation of groups of conditions, these groups do in fact pass the test set out by Mumford and Anjum—and do necessitate the manifestation of dispositions.

In the discussion above we looked at two possible situations: in one situation we have the group of conditions A, B, C and D; and in another situation we have A, B, C, and D but also the presence of an additional condition, E. We took it that in both situations we have the same group of conditions, A, B, C and D—and that in the latter case the group of conditions is present without the disposition manifesting, because E prevents this. But, on a different view of the individuation of groups of

conditions, we do not have the same group of conditions in both cases. When condition E is also present, we have a *new* group of conditions [A, B, C, D and E]. We do not have same group [A, B, C and D] with the addition of E.

On this view, it is true that the group of conditions [A, B, C and D] can necessitate the manifestation of the disposition—because whenever the group [A, B, C and D] is present, the disposition will manifest. There is no possibility of interference that would mean the group of conditions is present whilst the disposition fails to manifest. If there is interference due to another condition being present, the original group of conditions [A, B, C and D] is no longer present: instead we will have the different group of condition [A, B, C, D and E]. So, it is not the case that [A, B, C and D] could be present whilst the disposition fails to manifest.

But does it make sense to individuate groups of conditions in this way? I have used the example of a group of only four or five conditions to illustrate this point, but the groups of conditions that necessitate dispositions will in actuality be much larger than this. The way of individuating groups I have just introduced makes most sense when we are thinking about a totality of conditions capable of affecting whether an object manifests its disposition at a time—which in nearly all cases will include a vast number of conditions. Any attempt to circumscribe small groups of conditions from this totality at a time will appear ad-hoc. It is an entire state of affairs at a time that necessitates a disposition's manifestation. When we think of an entire state of affairs, a totality of conditions present at a time, it makes sense to think that if other conditions were added, we would be looking at a different state of affairs.

It appears, then, that whether dispositions are in fact necessitated to manifest by certain conditions turns on the way we individuate groups of conditions. If groups are individuated in the second way discussed above, Mumford and Anjum's argument fails: a disposition is necessitated to manifest by a totality of conditions at a time that brings about its manifestation.

So, to draw a conclusion about whether dispositions are necessitated to manifest requires that we look at the individuation of groups of conditions. But we do not need to go into this discussion of individuation here. This is because my interest is in our agential concepts rather than the nature of these powers in reality—my interest is in whether our *concepts* of disposition and ability involve necessity. In independent

reality it may be the case that given an entire state of affairs at a time, a disposition is necessitated to manifest. But it can be demonstrated that the way we think about stimulus conditions means that we do not conceive of dispositions as powers that are necessitated to manifest.

The objection to Mumford and Anjum's attempt to deny necessity relied on our considering the totality of conditions that contribute to the manifestation of a disposition at a time. When Mumford and Anjum only considered certain conditions in isolation, or small groups of conditions, they arrived at the conclusion that there is no necessary connection between stimulus conditions and the manifestation of a disposition. It was only when we looked at the idea of an entire state of affairs at a time that we opened up the possibility of dispositions' being necessitated to manifest. I will argue that in our thought about dispositions we conceive of stimulus conditions in isolation or in small groups, as Mumford and Anjum do, rather than having in mind the totality of conditions that could affect the manifestation of a disposition. This shows that, at least on our conception of disposition, stimulus conditions cannot necessitate the manifestation of dispositions.

To see that we normally think of stimulus conditions as isolated conditions, rather than the totality of conditions that could affect a manifestation at a time, we need only consider a few examples. Whether we gain our dispositional concepts through observation or—as Mumford and Anjum argue—through direct experience of our own agency, we only think of objects reacting to and being affected by a relatively small number of conditions. Paradigm examples of dispositions manifesting involve two things interacting: water boiling due to the heat of a flame; a window breaking due to a ball hitting it at a certain speed; salt dissolving because it is put in water; an agent being forced to lean forwards to make any progress walking against a strong wind.

In addition to the fact that paradigmatic cases of our thought about disposition involve a limited number of conditions, another reason to think we do not conceive of stimulus conditions as the totality of conditions that contribute to and prevent the manifestation of a disposition is that we cannot possibly be aware of all of the conditions that are present at a particular time, and the way in which they do in fact contribute to or prevent the manifestation of a disposition. As demonstrated by Mumford and Anjum, there are many different powers operating at any one time

that can interfere with, contribute to or prevent the manifestation of a disposition. The salt's dissolving in water depends on conditions such as the temperature being right and the water being unsaturated, but also on any number of other physical properties that do not enter into our everyday experience.

So, in everyday situations when we think about what contributes to a disposition's manifesting, we do not think in terms of a totality of conditions at a time. Rather, we think in terms of a few of the most relevant conditions. When we see salt being placed in water, it is entirely possible that it may fail to dissolve despite being in what we take to be the stimulus conditions for its disposition: it may be that there is another condition preventing the disposition from manifesting, such as the water's already being saturated. As far as we are concerned in this case, the disposition being in its stimulus conditions does not necessitate that it manifests.

We can accept all this whilst allowing that, in actuality, whatever total state of affairs actually does allow the manifestation or prevent it necessarily allows it or prevents it. It is simply the case that we do not conceive of stimulus conditions for a disposition as a totality of conditions that affect the manifestation of the disposition at a time. Thinking in terms of the total set of conditions that is actually affecting whether or not the manifestation takes place is much too complex and unnecessary for our getting by in our everyday interactions with other agents.

Although the above discussion in some ways concerns an epistemic point, my claim is not that because we do not think in terms of all of the conditions that contribute or prevent the manifestation of a disposition, we cannot know that powers will manifest in certain conditions. I can know that the sugar will dissolve if I put it in my tea, even if I also know that it need not dissolve just because I put it in my tea since something could happen to interfere with its dissolving. Similarly, I can know that my front door will open when I unlock it, even though I also know that it is possible for someone to have broken into my house and glued the door shut. My claim is just that, given the limitations in they way we think about stimulus conditions for dispositions, both a disposition manifesting *and* not manifesting is consistent with its being in the stimulus conditions we normally associate with dispositions. There is no case such that, given the presence of stimulus conditions of the kind we see as contributing to the manifestations of dispositions, the disposition

must manifest. Our concept of disposition is not the concept of a power that is necessitated to manifest by its stimulus conditions.

1.3 Necessity is not the difference between ability and disposition

We can now see that our concept of disposition is the concept of a power that is not necessitated to manifest even when in its stimulus conditions. What new light does this throw on the divide between dispositions and abilities?

Showing that dispositions need not manifest when they are in their stimulus conditions will never show that there is *no* difference between disposition and ability. It remains true that dispositions are more closely linked with stimulus conditions than the exercises of abilities: attributing a disposition tells us something about how an object is likely to behave in certain circumstances, whilst attributing an ability does not tell us anything about how an agent is likely to act. However, different dispositions are differently linked to stimulus conditions as well. Erasmus Mayr gives a taxonomy of different types of disposition, on which ascribing some dispositions tells us more about how an object that possesses them will act than ascribing others.[11] For example, an unstable explosive is often in conditions sufficient for the manifestation of its disposition to explode but may not in fact explode, whilst a piece of ice will usually melt when stimulus conditions for its disposition to melt. It seems that abilities are on this same spectrum—albeit far on the opposite end to powers such as the ice's disposition to melt.

Despite the remaining differences between ability and disposition, the finding that our concept of disposition does not involve necessity is important to our understanding of these two types of power. We can no longer entertain a picture like the one held by van Inwagen and Taylor, on which dispositions are necessitated to manifest by their stimulus conditions whilst abilities are never necessitated to manifest. Instead of this deep divide between the two types of power, we can see that their closeness to the conditions that allow their manifestation or exercise are simply on two ends of a scale. We must look elsewhere to find what it is that

11 (Mayr 2011) pp. 161-164

fundamentally separates the notion of the exercise of an ability from that of the manifestation of a disposition.

First, however, there is an issue raised by the view of dispositions we have been discussing which must be addressed in the remainder of this chapter. The idea that dispositions 'tend' or 'push' towards their manifestations when they are in their stimulus conditions appears to introduce a new difference between abilities and dispositions that might appear to be useful in better understanding the difference between an object's manifesting a disposition and an agent's exercising an ability. However, we will see that this is not an illuminating line of inquiry—the fact that dispositions 'push' towards their manifestations, when understood properly, does not point towards a difference between abilities and dispositions.

2.

Another Difference Between Ability and Disposition?

2.1 'Tending' and 'pushing'

This apparent difference between dispositions and abilities is best illustrated in a discussion by Markus Schrenk, who has a view of dispositions that employs an idea similar to Mumford and Anjum's 'tending towards a manifestation'.[12]

Schrenk argues that an account of dispositions must include the notion of a disposition aiming or pushing towards its manifestation. What highlights this need, for Schrenk, for a notion of 'pushing towards a manifestation' are 'antidote' cases.[13] Introduced by Alexander Bird in order to challenge the conditional analysis of dispositions, these are cases where a disposition is counteracted by something else so that the disposition fails to manifest even though it is in its stimulus conditions. A

12 (Schrenk 2009) pp. 156-157

13 Bird (1998) discusses these antidote cases. Mumford and Anjum (2011b) also use Bird's antidote cases in motivating the need for a dispositional modality.

deadly poison will not manifest the disposition to kill when ingested if an antidote is taken to counteract it. The poison does not *lose* the disposition, and it is in stimulus conditions for the disposition: the reason that it does not manifest its disposition is simply that the antidote is working against it.

Schrenk argues that a simple distinction between manifesting and not manifesting cannot adequately capture what is going on in these cases. Antidote cases show that there can be something in between a disposition's manifesting and not manifesting: the disposition is 'pushing, trying, or aiming to manifest.'[14] For example, consider an electron in an electric and a gravitational field. The electric field triggers the electron's disposition to move in one direction, and the gravitational field triggers the electron's disposition to move in the opposite direction. If the fields are of equal strength, the electron will remain stationary. However, this case is different from one where the electron remains stationary simply because it is not in stimulus conditions for its dispositions. When the fields are acting on it, its dispositions are being triggered but are being counteracted: there is a sense that the dispositions are active in some way even though they are not manifesting. Schrenk writes that 'antidote cases must differ from cases where the trigger is not at all pulled.'[15] To explain the difference, Schrenk suggests, we need the idea of 'pushing': even though they are not manifesting, in this antidote case the electron's dispositions are still 'pushing' towards manifesting.[16]

Mumford and Anjum are less explicit about whether they believe a distinction of the kind introduced by Schrenk is needed. But there is evidence that they, too, endorse a picture on which dispositions have three states: not manifesting; tending or pushing towards the manifestation; and manifesting. Mumford and Anjum contrast a case in which two teams in a tug-of-war simply hold the rope without pulling it, and a case in which both teams pull the rope but their strengths are equally matched, resulting in no movement.[17] They argue that even though there is no change brought about in either case, in the latter case there is causation whilst in the former case there is not. In the latter case, there are dispositions in stimulus conditions such that they could

14 (Schrenk 2009) p. 157

15 (Schrenk 2009) p. 159

16 Ibid.

17 (Mumford and Anjum 2011a) p. 30

manifest were they not being suppressed by other powers, but in the latter case these dispositions are not manifesting at all. In the former case they are tending towards a manifestation, but in the latter case they are not. So, like Schrenk, Mumford and Anjum appear to invoke a distinction between three states: not manifesting; tending or pushing towards a manifestation; and manifesting.

Both Mumford and Anjum's and Schrenk's views seem to include the thought that there are some cases where, although they are not manifesting, dispositions are somehow actively pushing towards an outcome due to being in their stimulus conditions. As Schrenk notes, we are restricted to using metaphors like 'pushing' that give agency to dispositions because the notion we are attempting to focus on is difficult to conceptualise further. But insofar as this notion of 'pushing' or 'tending' is part of the view of the nature of dispositions, it appears to introduce a difference between our concepts of disposition and ability.

On this view of dispositions these powers can be either inert, pushing towards their manifestation, or actually manifesting—but it does not seem possible that abilities could also be in all three of these states. We do not think of our abilities as pushing towards their exercises. Rather, they are either not being exercised or being exercised. For example, there are some situations in which conditions are right for the exercise of my ability to ride a bike—when I have access to a bike—and some situations in which conditions are such that my ability cannot be exercised—when I have no access to a bike. But the state of my ability is the same in both of these cases. It is not the case that in virtue of being in conditions that are right for its exercise, the ability tends or pushes towards being exercised—it is entirely up to me whether I exercise the ability. Whereas dispositions that are in their stimulus conditions push toward their manifestations even when they do not manifest, it is not the case that abilities push towards being exercised in virtue of being in conditions that are right for their exercise.

So, it seems that we have a picture on which dispositions can be in three states—not manifesting, pushing toward a manifestation, and manifesting—whilst abilities can only be in two—not being exercised, or being exercised. It may appear, then, that this is a core difference between ability and disposition on which we should be focusing our attention. We will see that this is not the case, however. We can

understand the idea of dispositions' pushing towards their manifestation in a way that shows this pushing is in fact a feature of abilities as well.

2.2 Pushing and manifestation

On the pictures presented by Mumford and Anjum and Schrenk, the disposition's pushing or tending is separate to its manifestation: they talk of dispositions pushing or tending *towards their manifestations*. I will propose that the notion of pushing does have its place in an account of disposition, but that rather than seeing dispositions as pushing towards manifestations, we should see that a disposition's pushing *is* its manifestation. The pushing is the disposition's bringing about a change, and not some preceding state where it is somehow active but not bringing about any change. Where the above accounts struggle to articulate the nature of disposition's pushing or tending, a view that identifies the pushing with the disposition's manifesting gives an overall more intelligible picture. Importantly, it also shows us that the fact that dispositions push or tend towards certain outcomes when in their stimulus conditions does not constitute a significant difference between dispositions and abilities after all. There is no third state between manifesting and not manifesting in the case of disposition.

Why do Mumford and Anjum and Schrenk hold the view that a disposition's pushing does not involve that disposition's bringing about any change? Their reasons for holding this view stem from the views they adopt of the nature of manifestations.

The main factor in Schrenk's holding this view is that he sees the manifestation of a disposition as an event, and in most cases as an achievement. What it is for the poison to manifest its disposition to kill me is for there to be an event of my being killed, and what it is for the salt to manifest its disposition to dissolve in water is for there to be an event of its being totally dissolved. Schrenk takes it that the manifestation of a disposition is that disposition's successfully, and fully, effecting the change that it is a disposition to bring about. There is no room in the manifestation itself for the disposition's pushing towards this change to be happening with the possibility of its being stopped or counteracted: once the

disposition has manifested the change has already happened. For Schrenk, a manifestation's occurring entails that the disposition has successfully brought something about. On this view, a disposition's bringing about change is all or nothing—either it successfully brings about a particular end state, or it does not effect any change at all.

Mumford and Anjum are committed to the view that a disposition's tending does not involve bringing about any change for a different reason to Schrenk. On Mumford and Anjum's account, the manifestation of a disposition is not an effect or event, but is simply the *contribution towards the effect* made by that particular power among the many that contribute towards any particular effect.[18] It is this that leads to the view that a disposition's tending does not involve it bringing about any change. Let us unpack Mumford and Anjum's notion of 'contribution' so we can better understand how this affects their conception of dispositions' pushing or tending towards their manifestations.

Mumford and Anjum make two claims that, taken together, commit them to holding that the manifestation of a disposition is a contribution to its effect. The first claim is that each disposition only has one manifestation type.[19] One advantage to this claim is that it allows a simple individuation criterion for dispositions: dispositions are individuated by what they are dispositions for. The second claim is that dispositions can bring about different effects depending on the other dispositions that they are manifesting with.[20] For example, a fire's disposition to heat things up can bring about water boiling, paper bursting into flame, or a painful burn depending on the other objects and powers that are present at the time. Mumford and Anjum hold a 'mutual manifestation' picture of powers, on which powers manifest only when they meet certain other powers: their 'mutual manifestation partners'. The effect that is brought about is then produced by these powers working together. This is a central part of Mumford and Anjum's view of dispositions, and is key to their picture of the way in which many powers can act together to bring about some outcome or counteract one another.

18 (Mumford & Anjum 2013) p. 556

19 (Mumford and Anjum 2011a) p. 7

20 Ibid. p. 27

These two claims commit Mumford and Anjum to seeing manifestations as *contributions* to effects, rather than effects themselves: they hold that each disposition can bring about many effects, so the claim that a disposition's manifestation is its effect would contradict their view that each disposition has only one manifestation type. Instead, their claim that manifestations are contributions allows them to say that each disposition has only one manifestation, even though, through their manifestations, dispositions can bring about different effects.

George Molnar gives an example that is helpful in illustrating the picture of dispositions on which their manifestations are contributions to effects.[21] A barge is being towed down a canal by two horses that are each on opposite sides of the canal. The ropes with which the horses are towing the barge are at 45 degrees to the barge. The manifestation of each horse's power to pull the barge is a force along the angled direction of the rope, but the effect is the straight-ahead movement of the barge. If either horse were acting alone there would be a different effect—the barge would be pulled towards one of the sides of the canal. The manifestations of the horses' powers, then, are different from the effects which they bring about. The manifestations are simply contributions to an overall effect.

This separation of contribution and manifestation shows why, on Mumford and Anjum's account, a disposition's tending or pushing cannot involve its bringing about change. If they hold that a disposition pushes or tends towards its manifestation, and its manifestation is a contribution to its effect, then its merely pushing or tending cannot involve any contribution to that effect at all.

A different view of manifestation—on which the manifestation of a disposition is neither a contribution nor an event, but is an effect and is a process—will throw a different light on these philosophers' reasons for seeing pushing in the way that they do. It will show that pushing is best understood as the disposition's manifesting and thereby bringing about change. This view will also diminish the difference between dispositions and abilities by showing that a disposition's pushing or tending is not an intermediate state between not manifesting and manifesting.

21 (Molnar 2003) pp. 194-196

2.3 The problems with seeing manifestations as contributions

My main motivation in replacing these views of manifestation with my own is to present an account of dispositions' pushing on which it is not some intermediate state between manifesting and not manifesting, and to thereby diminish this apparent difference between disposition and ability. But this is a very singular motivation, and may not seem enough to justify replacing Mumford and Anjum's carefully worked out view of manifestation. I will therefore spend some time showing that there are independent reasons for thinking that Mumford and Anjum's view of manifestation as a contribution should be rejected. In suggesting an alternative, I am not only serving the needs of my investigation into the differences between dispositions and abilities, but presenting a more plausible overall picture of the nature of dispositions.

There are two main reasons why Mumford and Anjum's view of manifestations should be rejected.

The first reason, highlighted in an argument by Jennifer McKitrick is that Mumford and Anjum's view leads to some implausible claims about manifestations.[22] If we accept that manifestations are contributions to effects rather than the effects themselves, then we must accept that a glass's breaking is not the manifestation of the glass's disposition to break when struck, or that the salt's dissolving in water is not the manifestation of the salt's solubility.[23] This seems contrary to the way we think about dispositions—to such an extent that it appears like a stipulation of a new meaning for the term 'manifestation'. Another counterintuitive consequence is that if manifestations are contributions to effects, then they are unobservable. On Mumford and Anjum's picture, we only observe the effects that manifestations contribute to, not the manifestations themselves. If theoretically necessary, perhaps these counterintuitive claims about the nature of disposition could be accepted. But they show us that the idea that manifestations are contributions should be avoided if possible.

22 (McKitrick 2010) pp. 80-81

23 Ibid.

A second argument against Mumford and Anjum's view of manifestations is that they introduce a new concept—a 'contribution'—which seems very difficult to clarify or increase our understanding of.

Mumford and Anjum introduce contributions so that they can claim that dispositions can bring about different types of effects depending on the situation even though each only has one manifestation type. But we are left with little idea of what it is about contributions that means the same contribution can lead to different effects, and how the contributions of different powers can come together to bring about different effects. The only similar idea we have is that of forces, which can come together and counteract each other to bring about different effects—and there is evidence that this is how Mumford and Anjum are thinking about the nature of 'contributions'. But, although this way of thinking works for some examples, it does not seem an appropriate way to think about dispositions in general.

Molnar's discussion of the two horses pulling a barge is what Mumford and Anjum point to in explaining how they understand the manifestations of dispositions.[24] It is telling that the example of the two horses sees the manifestations—their 'contributions'—as forces acting in different directions, which can combine in order to produce movement in a single direction. Although the notion of contribution seems to work in the case where we see the manifestations as forces, this is not how it works in all cases. For example, what is the contribution made by the fragile glass to the overall situation when it manifests its disposition to break when struck? It is not the breaking itself, and it does not seem that the contribution can be a force. How does the glass' contribution work with the contribution of the hammer that strikes the glass? Thinking about forces does not give us any help here, or in many other similar cases of interaction between two objects. Without some clearer idea of what 'contributions' are meant to be, it does not seem that they can play the role they are intended to. Mumford and Anjum seem to think of these contributions as analogous to forces, but that only seems appropriate in particular cases and not in all cases of dispositions manifesting.

So, there are independent reasons to try to find some other way to view manifestations of dispositions, that will help us accommodate the idea of

[24] (Mumford & Anjum 2013) p. 556

disposition's pushing without seeing this as some third state between manifesting and not manifesting.

3.

Manifestations as Effects

3.1 Processes

In this section I will present a view on which the manifestations of powers are the effects produced by those powers. Many philosophers who hold that manifestations of powers are effects see these effects as events—as the event of the glass breaking or the event of the salt dissolving. I will argue that in order to retain Mumford and Anjum's and Schrenk's insight about the 'pushing' that is central to our notion of power whilst at the same time accepting that manifestations are effects, we must see the manifestations of powers as processes which persist through time.

I will begin by illustrating the distinctive features of processes, drawing on Rowland Stout's arguments for a distinction between events and processes.[25] There is controversy over how processes should be understood, and I do not claim that Stout's overall account is the correct one. I only claim that processes have the core features mentioned below. These features are enough for the claim that manifestations are processes to be a plausible and helpful one.

Stout's aim is to draw out a distinction between events and processes that he believes is already present in the way we think about things going on. His initial way into the distinction is through grammar. He writes: 'When something is/was/will be/etc. happening, let me call what is/was/will be happening a *process*. When something happened/will happen/etc., let me call what happened/will happen an event.'[26] So when we say something like 'the salt is dissolving', we are talking about a

25 (Stout 1997); (Stout 2007)

26 (Stout 1997) p. 19

process of the salt dissolving. When we say something like 'the salt dissolved', we are talking about an event of the salt dissolving.

With this initial separation in place, we can think about the ways that processes are different from events. The core difference between the two is that whilst events extend through time, processes persist through time analogously to the way that objects persist through time. Stout claims that 'there is something absurd about saying that at any one time while something is happening only part of what is happening is present'—just as it is absurd to say that an object is not wholly present at any one moment at which it exists.[27] At a moment in time when a glass is breaking, it is the process of the glass breaking that is happening, not just a part of the process. This contrasts with events, since at any moment in time it is only a part of the event that is present.

The idea that processes persist through time whilst events do not forms the basis of Stout's argument that processes are not identical to events. Since processes persist through time, their identity does not depend on their duration, just as an object's identity does not depend on how long it exists. The process of a chicken walking across the road would have been the same token process whether it ended with the chicken making it across the road or it ended some time earlier with the chicken being hit by a bus. In contrast, the identity of an event does depend on its duration. The event of the chicken walking across the road would not have been the same token event if it had taken longer for the chicken to reach the other side or if the chicken had only made it halfway across the road. Given that events and processes have different identity conditions, Stout concludes that a process cannot be identical to an event.

Having established this difference in identity conditions between events and processes, we can now look at why understanding manifestations as processes helps us accommodate the idea of dispositions tending or pushing.

27 (Stout 1997) p. 25

3.2 Manifestations as processes

In Stout's discussion of the identity conditions of events, we saw that an event could not have been interrupted before its actual end point—the duration of a token event is essential to it. Events are not the kinds of things that can unfold or be interrupted. They simply extend through time. Stout's grammatical distinction further underlines the fact that we do not think of events as things that can be interrupted. An event is something that has happened or will happen, not something that is happening. Once there is an event on the scene, such as the event of a glass breaking, then there is nothing can be done to stop the glass from breaking—no intervention is possible because the glass's breaking has already occurred.

This shows, as we saw from Schrenk's account, why thinking of manifestations as events leads to looking for the pushing that characterizes disposition *before* the power manifests. If the manifestation is an event, once the power manifests there is no possibility of interruption or interference that could stop the object with the power doing what it aims towards. On this view, once the water manifests its power to dissolve the salt—in other words, once an event of the salt dissolving takes place—there is no possibility that anything could interfere with the salt dissolving. So, the 'pushing' towards some outcome, where there is always the possibility that the outcome could fail to occur because of counteraction by other powers, can only be located before the power manifests.

If, on the other hand, the manifestation of a power is a process, then there is no need to restrict the pushing to the time before the disposition starts to manifest. Even once a process has begun there is the possibility that it can be interrupted, so the fact that the water has started to manifest its disposition to dissolve the salt does not necessitate that the salt fully dissolves. A power can manifest without bringing about the outcome it aims towards, since the manifestation can be interrupted once it has begun. On this view, it makes sense to locate the pushing during the manifestation of the power. When the manifestation is occurring, the power is pushing towards some outcome and there is still the possibility that this outcome will not occur due to the manifestation being interfered with or interrupted.

Adopting this view allows us to capture the notion of pushing or tending discussed by Mumford and Anjum and Schrenk, but it requires a change in Mumford and

Anjum's central contention that 'Dispositions only tend towards their manifestations; they do not necessitate them'.[28] With this new picture of the nature of 'pushing' as the manifestation of the disposition, we can no longer say that dispositions tend or push towards their manifestations. We must now say that dispositions tend towards bringing about certain situations. When in its stimulus conditions, a disposition will push towards a certain state of affairs, but the process whose end point is this state of affairs may be interrupted or prevented from progressing. It remains true that dispositions are never necessitated by their stimulus conditions to manifest or to bring about a certain effect.

Let us look more closely at how this conception of a manifestation as a process allows us to capture the notion of 'pushing' as simply an aspect of the manifestation itself. Recall Schrenk's initial characterization of pushing: Schrenk distinguishes between three states: a disposition's being inert, a disposition's manifesting and a disposition's pushing towards its manifestation. My view is that dispositions are either manifesting or not manifesting, and that this is sufficient to account for the idea that dispositions 'push' towards certain outcomes. A dispositions' pushing towards its outcome just is its manifesting.

Take the case of a poison being administered, and its antidote being administered directly afterwards, preventing me from being harmed by the poison. One way to understand what happens here is that the antidote prevents the poison from manifesting its disposition to harm me. But it is clear that this is very different from a case in which the poison is not administered at all, and so does not manifest its disposition. For this reason Schrenk would say that here the poison is 'trying' to manifest but being prevented from doing so by the antidote.

On the view of manifestation as a process, we can see that in this case the poison does manifest its disposition to harm me, but that its manifestation is interrupted before it reaches the end point of harming me. The poison is acting on my body, and after the antidote is administered will react with the antidote. But because of the antidote's work, the process the poison is taking part in will not continue to the point where I am harmed. It is in this way that the poison is 'pushing' towards the outcome of my being harmed—not because it is in an intermediate state between

[28] (Mumford & Anjum 2011b) p. 380

manifesting and not manifesting, but because it is manifesting its disposition in a process towards my being harmed.

Schrenk cannot allow for this, and finds himself having to introduce the idea of pushing but not manifesting, because on his view the manifestation of the poison's disposition is an event: the event of my being harmed. If the end point of my being harmed is not reached, the event of my being harmed never happened, and so no manifestation has taken place. Schrenk has no way to articulate the idea that the disposition could be manifesting even if it is interrupted before it reaches this end point. For Schrenk, either the end point is reached or there is no manifestation at all. The view of manifestation as process allows us to account for the idea that the poison is doing something even if it is ultimately prevented from reaching the end point of the process.

The view of manifestations as processes also helps us account for 'pushing' in other types of case that may initially seem to require an intermediate state between not manifesting and manifesting. This is because, in many of these cases, there is a process occurring with which we can identify the pushing. Consider the example of two teams taking part in a tug of war, whose strength is balanced so that neither team is actually moving. Schrenk would say that the participant's powers to move the rope are not manifesting, since the rope is not moving—their powers are merely 'pushing' towards moving the rope but being counteracted. But in any case of two teams pulling a rope there is a process occurring—muscles tensing, leaning, the rope straining and its fibers becoming stretched. Due to the perfectly balanced influences on the rope this process does not involve the rope moving a great deal in one way or another.

My claim is not that all 'antidote' cases involve the participants actually manifesting their dispositions. My claim is simply that in all cases where there is pushing, this is due to a process taking place. There may be some cases where the disposition is being prevented from manifesting at all, and so there is no pushing occurring. In the example of the electron stationary between two magnetic fields, if there is truly no movement of the electron at all we might have to say that there is no manifestation: in the absence of movement it is difficult to justify saying there is a process taking place. But in the case of no manifestation taking place, we should also say that there is no 'pushing' towards an outcome. If we did want to say that the electron were

pushing towards these outcomes, we must find some way to say that there is a disposition being manifested.

An advantage of seeing pushing as taking place as the power is manifesting is that pushing becomes less mysterious, since we can see it as something that actually happens. On Mumford and Anjum's view, powers push towards their manifestations—but it is not clear what contribution this pushing towards a manifestation can make to what happens in a situation, since the pushing must occur before the power actually makes a contribution.

On the view I have proposed, powers push towards certain outcomes, and the pushing takes place when the manifestation is occurring: the flame's power to burn the paper pushes towards the paper being fully consumed by the flame, and the flame is pushing towards consuming the paper as it is in the process of consuming the paper. On this view, pushing is not mysterious since it is simply activity on the part of the object whose disposition is manifesting.

There is a final clarification to be made about the view of manifestations I have presented. It is important to note that not all manifestation are processes—this is because not all manifestations involve something happening. Some manifestations are simply states of objects. For example, a magnet that is stuck to a fridge is manifesting its disposition to be attracted to certain objects. But there is no change occurring and no process going on. The manifestation is simply the magnet's being stuck to the fridge.

In many cases where we think of dispositions manifesting we think of objects interacting and undergoing change. But some cases are like that of the magnet: there is a disposition manifesting but no change taking place as it does so. I do not take this to present a problem for my proposal, since when a manifestation is a state the manifestation is still an effect, rather than a contribution to an effect. Further, we must admit these cases whether we see the manifestations that do involve change as events or as processes—such cases are not an exception only to the view I have proposed. Finally, in cases like those of the magnet, there is no pushing on the part of the disposition but only manifestation. These cases are consistent with my view that a dispositions' pushing is a matter of its manifesting in a process that can be altered or prevented from reaching its outcome.

3.3 Dispositions and abilities

When we understand pushing as taking place as a disposition is manifesting, we can see that the idea of pushing does not introduce a deep difference between dispositions and abilities. Let us recall how the idea of 'pushing' initially threatened to introduce a difference between dispositions and abilities: on Mumford and Anjum's and Schrenk's views, dispositions may be in a state between manifesting and not manifesting, in which the disposition is somehow pushing towards its manifestation but not yet manifesting. This third state of pushing in between manifesting and not manifesting is clearly absent from cases of abilities, which are only ever either being exercised or not being exercised. Even when in conditions favourable to their exercise, abilities do not 'push' towards their being exercised—they are either being exercised or not.

On my view that the pushing is the manifestation of the disposition, there is no third state in the case of dispositions that is absent in the case of ability. Both dispositions and abilities are either manifesting/being exercised, or not manifesting/not being exercised. The disposition's pushing is simply its taking part in a process of change that could be interrupted or altered before it reaches its end point. The same occurs in the case of an ability being exercised: there is a process of change that could be interrupted or altered. My exercising my ability to open a door is a process which could be interrupted if someone starts pulling back on the other side of the door, or prevented if the door is glued shut—just as the poison's manifesting its disposition to kill me is a process that could be interrupted if the antidote is administered after the poison, or prevented if the person to whom the poison is administered is immune.

We have seen that pushing actually refers to a feature of dispositions that is also a feature of abilities: their bringing about a process of change that can be interrupted before reaching its conclusion. But what should we conclude from this about the notion of pushing—does this show that our notion of pushing should also apply to abilities, or that it should apply to neither abilities nor dispositions? We should draw neither of these conclusions. The aim here is not to completely oppose the intuition that dispositions push towards some end point. It may still be appropriate to think of dispositions pushing towards some end point whilst abilities do not. It is just that

this difference is not based in there being some third state dispositions can be in that abilities can not.

Our intuition that dispositions 'push' but abilities do not isn't due to the existence of a state between manifesting and not manifesting in the case of dispositions, and the absence of this state in abilities. Rather, it is due to the simple fact that we think of dispositions as having much more specific stimulus conditions than abilities, and of the occurrence of a dispositions' manifestation as having a much tighter connection with these conditions. Salt will only dissolve in very specific circumstances—if it comes into contact with water that is not too cold, that is not saturated etc.—and, despite the fact that it is never necessitated to dissolve it will *usually* dissolve if it is put in these circumstances. Our abilities often have very wide stimulus conditions and their exercise is not closely connected to these conditions. The stimulus conditions for my ability to raise my arm are simply any circumstances in which I am not restricted, but when these circumstances are present I will not usually raise my arm. Our intuition that dispositions push towards some effect when they are in stimulus conditions is due to the fact that, usually, when they are in stimulus conditions, dispositions will manifest. There is no similar link with stimulus conditions in the case of abilities, which is why we do not have this intuition that abilities are reaching towards particular outcomes just because certain conditions are present.

3.4 How to individuate dispositions

There remains one loose end to tie up. Mumford and Anjum's main reason for claiming that manifestations must be contributions is that they believe it is the only conception of manifestations that gives a plausible view of the individuation of dispositions: it allows them to hold the same disposition may produce different effects when paired with different manifestation partners, but also to hold that each disposition only has one manifestation—making it easier to individuate dispositions by their manifestations. On the view of manifestations I have proposed, the manifestation is the effect: when two dispositions manifest together the manifestation of both dispositions is the same process.

For my proposed view of manifestations to be acceptable, I must show how it avoids problems about the individuation of powers, thereby demonstrating that the above does not give a reason to prefer Mumford and Anjum's view of manifestations over the one I have proposed.

First, let us remind ourselves of Mumford and Anjum's position on the individuation of dispositions. There are three theses held by Mumford and Anjum that lead them to believe we must understand manifestations as contributions in order to have a plausible view of the individuation of dispositions:

1) Powers have different effects with different mutual manifestation partners.
2) Two or more separate powers have the same effect when they mutually manifest.[29]
3) The manifestation type determines the identity of the power.[30]

1) and 2) are part of their conception of mutual manifestation. 3) they appear to take to be common-sense, since we normally identify powers by what they are powers *for*. As mentioned above, Mumford and Anjum claim that the manifestation of a disposition must be its contribution to an effect rather than the effect itself due to the conflicts that would occur between 3) and the other two theses if manifestations were taken to be effects. If the manifestation were an effect, it could not determine the identity of a disposition if the disposition had different manifestations in different situations. Nor could the manifestation determine the identity of the disposition if two separate dispositions could produce the same effect.

In this section we will see that there are at least two plausible alternative ways to think about the individuation of powers that are compatible with my view of manifestations as processes. These alternatives require taking another look at Mumford and Anjum's three theses, but they allow for a plausible view of the individuation of dispositions without straying too far from Mumford and Anjum's picture. It is not important for my overall argument which of these alternatives is preferred—all that matters is that my view of manifestations as processes does not lead to implausible consequences for the individuation of dispositions.

29 (Mumford & Anjum 2011a) p. 27

30 Ibid. p. 7

The first alternative involves taking another look at thesis 1). Mumford and Anjum argue that the same disposition has different effects with different manifestation partners, leading them to say that a disposition cannot be individuated by its effect. What they have in mind are dispositions like magnetism—a piece of metal that has the disposition of magnetism may manifest this by attracting a piece of iron towards it, but in other circumstances where there are countervailing dispositions may only succeed in making the iron vibrate. It cannot, then, be the effect that is the manifestation of the disposition and through which one picks out the disposition. Mumford and Anjum's thought is that instead there must be something else underlying these effects: the contribution that the disposition makes in producing each of these different effects.

We can avoid this problem by claiming that there is not one disposition—such as magnetism—but a variety of other dispositions with more restricted effects—such as the disposition to attract iron, the disposition to make iron vibrate, and so on. In doing so, we are effectively rejecting 1). We are claiming there are no dispositions with multiple different effects—instead, there are many different dispositions, each with only one effect. This strategy involves increasing the number of dispositions there are in the world by a great deal, but it allows us to see how dispositions can be individuated by their manifestations even if their manifestations are their effects.

It is not immediately clear that multiplying dispositions in this way solves all of Mumford and Anjum's problems with individuating dispositions when manifestations are effects. Thesis 2) may still appear to present a problem. This thesis states that when two dispositions mutually manifest—the disposition of the salt to dissolve in water and the disposition of the water to dissolve the salt—there is only one effect. There is only one process of the salt dissolving in/being dissolved by the water. How, then do we distinguish the water's disposition to dissolve the salt and the salt's disposition to be dissolved by the water, as they both have the same manifestation?

It seems that the only way to avoid the problem is to change the way we understand 2) as well. We can note that although the salt and the water are participating in the same process, both are doing very different things. The salt is dissolving, whilst the water is changing its make-up to become more saline. So, although the salt's disposition to be dissolved and the water's disposition to dissolve the salt ultimately

have as their effect the same token process, they do this through taking part in two separate processes that make up the shared process.

Although this proliferation of dispositions, each with only one effect, gives us a clear way to individuate these powers, some may find themselves unable to accept that we should think of dispositions in this way. The second alternative view of the individuation of dispositions does not require that we multiply dispositions like this—and is in my view the more plausible option of the two.

The second alternative focuses on thesis 3). Mumford and Anjum hold that the manifestation of a disposition determines its identity—but seem to rule out that a disposition can have more than one manifestation. However, many philosophers propose views on which dispositions are 'multi-track'—the same disposition can have different manifestation types.[31] The idea that we identify dispositions by what they are dispositions for does not itself prevent us from seeing dispositions as being multi-track. A multi-track dispositions view better allows us to see how a disposition can have different manifestations with different manifestation partners. It also allows us to see how a disposition can have the same manifestation as its mutual manifestation partner—we still see the two as separate powers because, although this particular manifestation type is shared by the two dispositions, each disposition has different manifestation types not shared by the other.

Mumford and Anjum do not consider this option when discussing the individuation of dispositions. There are two possible reasons for this. One possible reason is that they think the claim that there are multi-track dispositions (dispositions with more than one manifestation) is unmotivated. Another possible reason is that they think that the notion of multi-track dispositions causes even more problems when it comes to individuating dispositions. We will see that neither of these reasons are good ones, and so individuating dispositions by groups of manifestations remains an option.

First, let us look at the idea that proposing multi-track dispositions is unmotivated. Some may argue that the onus is on the proponent of multi-track dispositions to say why we should accept that there are fewer dispositions with multiple manifestations,

31 Gilbert Ryle (1949, p. 114) gives an early discussion of multi-track dispositions. Other, more recent accounts include (Cartwright 1998), (Bird 2007) and (Vetter 2013).

rather than many different dispositions with single manifestations. This task has already been completed by the previous discussion. If one believes that proposing the existence of many different dispositions—one disposition for every manifestation—is more plausible, then the first view discussed in this section should be taken. This multi-track view is an option for those who think that it is implausible to propose so many different dispositions. The only other alternative to multi-track dispositions—the only way to hold that dispositions have only one manifestation without accepting the existence of many different dispositions—is to take the view that manifestations are something like Mumford and Anjum's 'contributions to effects'. We have seen why this alternative is not attractive. So, there is motivation for accepting a multi-track dispositions view.

Now let us look at the more serious objection, that multi-track dispositions cause further problems when it comes to individuating dispositions. One example of this criticism is presented by E. J. Lowe:

> *Once we allow that powers may genuinely have multiple manifestation-types which don't fall under any unified description, it becomes unclear why we should think that a single object may have many different powers rather than just one—a power to do all the things that it can do. And that would render the notion of power a rather feeble and trivial one.*[32]

Lowe's worry is that the notion of a disposition with more than one manifestation will ultimately lead us to the view that each object only has one disposition—since if we accept that more than one manifestation can belong to the same disposition, there is no reason why we should not just say that every manifestation belongs to the same disposition. But all we need to avoid Lowe's problem is some principle to guide us in grouping together manifestations. And such a principle is readily available.

Neil Williams, in a paper defending the notion of multi-track dispositions, discusses what he terms 'clustering':

32 (Lowe 2010) pp. 11-12, quoted in (Williams 2011)

For event types φ and ϕ, where φ-ing ≠ ϕ-ing, the power to ϕ and the power to φ are clustered just in case the class of objects with the power to ϕ is identical with the class of objects with the power to φ.[33]

For example, the set of objects with the disposition to attract certain other types of object is identical with the set of objects with the disposition to repel certain other types of object—these dispositions are always found together in objects that are magnetic. We would then say that these dispositions are 'clustered'. This notion of clustering provides us with a principle to guide our grouping together of manifestations.

Williams' principle is not suggested as a hard and fast guide to individuating dispositions, nor is it suggested as a necessary or sufficient condition for the disposition to φ being identical to the disposition to ϕ. Rather, it is suggested as a consideration that may incline us to believe that two different manifestation types may be manifestations of the same disposition, all things considered. This combats Lowe's claim that on a multi-track picture there would be nothing to guide us in individuating dispositions.

Indeed, Williams' principle alone is cannot guide us in individuating dispositions, since we also need a principle for deciding how to individuate types of process before we can start deciding whether two different process types are the manifestations of the same disposition. For example, is the elastic band's stretching 1.5cm the same process type as the elastic band's stretching 1.6cm? It seems natural to say that they are the same process type—but the important point to note is that this is a question we must ask no matter whether we accept single-track or multi-track dispositions. It is not a problem created by adopting a multi-track view.[34]

---

33 (Williams 2011) p. 582

34 It is worth noting that the principle adopted for individuating process types can impact on how plausible Williams' 'clustering' principle is. If we decide that the elastic band's stretching 1.5cm and 1.6cm are different types of process, we may yet want to say that both of these processes are manifestations of multi-track dispositions. Yet we could not use Williams' principle as a reason to say this. There are surely some elastic bands that can stretch 1.5cm but not 1.6cm, whilst others cannot. So the class of objects with the disposition to stretch 1.5cm is not identical with the class of objects with the disposition to stretch 1.6cm. However, the problem here may be a view of process individuation that is too fine-grained, rather than an issue with Williams' general principle—which is not intended to apply in all cases, but is intended to act as a guide. This issue is discussed further in Chapter 5.

What looking at a principle like Williams' shows is that we are not in the situation that Lowe describes, where accepting that dispositions are multi-track leads to identifying all dispositions with one another and seeing all manifestations as manifestations of the same disposition. There are principled ways of deciding when manifestations belong to the same disposition and when they do not.

The two alternatives I have discussed in this section show that we can view the manifestations of dispositions as processes without running into problems over how dispositions are to be individuated. A plausible approach to individuating dispositions does not require viewing their manifestations as contributions to effects.

Conclusion

We have now established that our concept of disposition is not the concept of a power whose manifestation is necessitated by stimulus conditions. Rather, we conceive of and experience a world in which it is always possible that a disposition might not manifest even if its stimulus conditions are present. This eliminates the possibility that what makes the difference between dispositions' manifesting and abilities' being exercised is a difference in necessity. It is not the case that inanimate objects are necessitated to manifest their dispositions in some circumstances, whilst human agents are never necessitated to exercise their abilities. Neither dispositions nor abilities necessarily manifest in any circumstances.

As discussed above, it is clear that there are still distinctions between dispositions and abilities in terms of how closely connected their manifestations or exercises are to their stimulus conditions. This is what accounts for our intuition that dispositions 'push' towards certain outcomes while abilities do not. We have seen that dispositions' pushing is in fact nothing more than the disposition bringing about a process that could be interrupted. Dispositions do not differ in this respect from abilities.

We have spent time looking at some apparent features of disposition and ability that may be thought to introduce an important difference between the two, and have found that these do not aid us in illuminating the distinction between an agent's

manifesting a disposition and exercising an ability. In the next two chapters I move onto looking at some features of ability that do constitute important differences between the two types of power: abilities are possessed and exercised by embodied agents, and involve agents being in control of their bodies.

Chapter 4

Self-movement and Bodily Control

We have now established that our conception of a disposition is the conception of a power that is never necessitated to manifest. Like the exercises of abilities, the manifestations of dispositions are never necessitated by the presence of particular circumstances—even if being in particular circumstances makes it much more likely that an object will manifest its disposition.

The key difference between abilities and dispositions lies elsewhere. What is unique about abilities is not that their exercises are not necessitated, but that the exercise of an ability involves self-movement on the part of the agent. What this means is that the agent exercising an ability is in charge of what she does and how she moves in a way that an agent manifesting a disposition is not. In this chapter and the next we will look at how best we can draw out and understand this feature of abilities.

The first step is to clarify what is meant by 'self-movement', and look at the role of self-movement in distinguishing ability from disposition. We can then turn to considering how best to understand this feature of ability. We will look at attempts to provide accounts of self-movement by Harry Frankfurt and Helen Steward, who both take very different approaches to understanding what it is for a human agent to be in charge of how she moves. Despite their different approaches, both philosophers' views of self-movement are subject to the same problem. This will allow us to see the mistake we must avoid in giving an account of self-movement and its role in ability.

1.

Self-movement

1.1 What is self-movement?

As with many terms in philosophy, 'self-movement' is used by different philosophers to draw different distinctions. It is therefore important to make clear exactly what is meant by the term as it will be used in the rest of this chapter. Some may have a very inclusive view of what it is to be self-mover, taking being a self-mover to come along with simply being an animate agent. In the sense that I will be using the term, self-movement requires more than being an animate agent. It requires that the agent is in charge of how its body moves. To illustrate the distinction, let us look at these two different uses of 'self-movement' in more detail.

On the more inclusive view, self-movement is movement of an agent that is brought about by something internal to an agent, rather than by something external to the agent. Fred Dretske seems to be thinking of self-movement in this sense when he argues that behaviour is 'movement that has its causal origin within the system whose parts are moving.'[1] Self-movement, on this view, is 'internally produced movement', and self-movers are those things whose movements are at least sometimes brought about from within.

In this sense of 'self-movement', inanimate objects like footballs, pebbles, or road signs are excluded from the category of self-movers: these objects do not move unless something else causes them to move. But plants, robots and other kinds of machinery should be counted as self-movers in this sense. When a flower grows and twists to face the sun, it seems that there is nothing external causing it to do this—rather, its movements result from the activity of internal mechanisms. Similarly, robots and machines that are programmed to move in certain ways—such as the robotic arms used on assembly lines—engage in movements that are the result of internal mechanisms. Also swelling the ranks of self-movers on this understanding

1 (Dretske 1988) p. 2

of 'self-movement' are nearly all animals. This includes humans, but also includes living things like dust mites or some single celled organisms, which also get around under their own steam. On this understanding of the term, 'self-movement' is taken to mean that the organism moves *by itself*. To be a self-mover is simply to be animate.

'Self-movement', understood in the sense I will be interested in, requires more than an agent's movement being 'internally produced'. It requires that the agent is in charge of the movements of its body—in charge of how and when its body moves. It is not enough that the movements of the agent are brought about by some mechanism internal to the agent or by some part of the agent. The agent's movements must be brought about by the agent. On this understanding of self-movement, being a self-mover is not simply a matter of an agent moving *by itself*—rather, it is a matter of the *agent moving itself*.

Where we cannot see an agent in charge of the movements of its body, we cannot see self-movement in this sense taking place. So, although plants and machines move by themselves, we cannot see them as self-movers in this sense. The flower's movements of turning to face the sun are the result of the mechanisms internal to the plant reacting to factors such as the direction and intensity of light. So, it is mechanisms within the plant, in combination with external conditions, that are causing the plant to move—the plant itself is not in charge of its own movements. Many other organisms are excluded from the category of self-movers when we use 'self-movement' in this sense. Single-celled organisms and dust mites may move by themselves, but they, too, do not appear to be in charge of their own movements. Humans are certainly self-movers, and so are many non-human animals, but the label 'self-mover' is a much more exclusive one on this understanding of the term.[2]

We take it that abilities are exercised at will, whilst dispositions manifest in response to stimulus conditions. As we have seen, though, what distinguishes these two notions is not that the exercises of abilities are not necessitated. Neither the exercises of abilities nor the manifestations of dispositions are necessitated, and so

2 Of course, coming to a conclusion about which agents are and are not self-movers is not as simple a matter as is presented here. In Chapter 6 we will look at this question in more depth. At present, my aim is simply to introduce the distinction between these two senses of 'self-movement' in order to clarify the way I am using the term in the following discussions.

the absence of necessity cannot be all there is to an ability's being exercised at will. Rather, the idea that agents exercise their powers at will involves the idea that the agent is in charge of how it moves: the agent is a self-mover, in charge of its own movements. An agent that is not a self-mover—an agent whose movements are brought about by other agents or by internal mechanisms—cannot possess or exercise abilities. If the agent is not in charge of the movements that it is making, it must be the case that it is manifesting a disposition rather than exercising an ability.

We have now seen something of the notion of self-movement, which has been explicated in terms of the agent being in charge, or in control, of the movements of its body. The key question we must now ask is what exactly it is for an agent to be in charge of its the movements of its body. This discussion will further demonstrate the close connection between our notions of self-movement and ability. Those who attempt to give accounts of self-movement and an agent's being in charge of its body in terms of notions other than ability end up presenting a picture on which agents are separated from their bodies.

1.2 Philosophers' views of self-movement

What is it for the agent to be a self-mover—to be in charge, or in control, of the movements of her body?[3] A picture that no contemporary philosopher wants to endorse is one on which the agent is a homunculus in the mental 'control room' of the body, and moves her body as if she is operating a puppet or a machine. This picture is so unappealing because it separates the agent from her body. On this picture it seems that the agent is no longer a self-mover at all, since the body becomes something separate from the agent that she acts on as if she were acting on another object.

Unfortunately, many attempts by philosophers to give accounts of an agent's being in control of her body come dangerously close to giving us this kind of picture. The agent, on these views, relates to her body as 'other'—the agent merely causally

3 I am taking 'in charge of one's body' and 'in control of one's body' to have the same meaning. In the following discussion I will most often use the term 'control' to refer to this idea, as this is the more commonly-used term.

affects movements of the body, rather than moving the body itself.

I will begin by looking at Harry Frankfurt's influential discussion of the notion of 'guidance', which can be read as an attempt to provide an account of the agent's control over her body in acting.[4] I will argue that his view that the agent guides her actions leads to the separation of agent and body. Frankfurt's starting point is the widespread reductive assumption that actions are essentially physical bodily movement events. It is not surprising, then, that he has this problematic view of the agent's being in charge of, or in control of, what she does. It *is* surprising that this problematic view has found its way even into recent antireductive accounts of action, such as the account given by Helen Steward.[5] What makes this even more surprising is that Steward's overall project is to develop an account where agent and body are seen as a whole integrated system.[6]

Although these philosophers' accounts ultimately provide little positive help in understanding the notion of self-movement, it will nevertheless be instructive to examine their views. Given that the problem of alienating the agent from the body appears to crop up in very different approaches to agency, it is important to diagnose why the problem arises in these very different views, in order to see how we can avoid it. Our discussion will show us that there is in fact little we can say about what it is for the agent to be a self-mover, in charge of the movements of her body, that does not lead to the separation of agent and body. The best way to explicate the notion of self-movement is in terms of agents exercising their abilities.

4 (Frankfurt 1978)

5 (Steward 2012)

6 Ibid. p. 49

2.

Frankfurt's Account of 'Guidance'

2.1 Introducing 'guidance'

Frankfurt is not specifically concerned with self-movement—his aim is simply to 'explicate the contrast between what an agent does and what merely happens to him.'[7] However, it seems that his notion of 'guidance' is intended to explain what I have referred to as the agent's being in control, or in charge, of the movements of her body. Looking at Frankfurt's motivation for giving his account will demonstrate this.

Frankfurt introduces guidance into his view of action so that he can avoid a problem he finds with the standard causal theory of action (CTA). The standard CTA account takes actions to be bodily movements that are caused by the agent's mental states. On this account, my action of raising my arm is an event of my arm going up that has been caused by a rationalizing belief and desire. The only difference, then, between an action of raising my arm and a non-action movement of my arm going up is in the causal history of these events. The action is a movement of my arm caused by certain mental states, whilst the mere movement is a movement of my arm caused by something else, such as a spontaneous muscle spasm.

Frankfurt's problem with this view is that the only difference between an action and a mere movement comes at the beginning of the event, whilst what is happening after the initiation of the event is the same in both cases. This means that the agent's role in raising her arm only comes at the beginning of the movement: during the time that the arm is rising the agent is not involved. The standard CTA view 'require[s] nothing of an agent, once the specified causal antecedents of his performing an action have occurred, except that his body move as their effect'.[8] Frankfurt argues that we need an alternative picture to account for the fact that 'during the time a person is performing an action he is necessarily in touch with the

7 (Frankfurt 1978) p. 157

8 Ibid.

movements of his body in a certain way.'[9]

We can see, then, that Frankfurt's target is similar to ours. He states his aim in general terms—to discover the difference between an action and a mere bodily movement—but it appears that the aspect of action he focuses on is the agent's being a self-mover, in control of the movements of her own body throughout the time that she is acting. Frankfurt introduces the notion of 'guidance' to account for the agent's control of her body during her actions.

Frankfurt's 'guidance' is a generic notion that has a place both in cases of human agency and agency of other kinds. I will first explain the notion generally, before going on to look at its application in the specific case of human agency.

Frankfurt writes that a movement is being guided when 'its course is subject to adjustments which compensate for the effects of forces which would otherwise interfere with the course of the behaviour, and when the occurrence of these adjustments is not explainable by what explains the state of affairs that elicits them.'[10] This explanation shows that Frankfurt conceives of guidance and the purposiveness of a movement as very closely linked: a movement is guided when there is some mechanism ensuring it reaches some particular end state, which can be seen as its goal. An example given by Frankfurt of a movement that is guided is the dilation of pupils in response to a change in light.[11] The purpose of the pupils' movement is to reach a size that allows the right amount of light into the eye. If there is a sudden change in light, a mechanism will ensure that the goal is still achieved by adjusting how far the pupils dilate. Frankfurt argues that a movement's being under the guidance of a mechanism does not require that the mechanism actually adjusts the movement. Perhaps there is no sudden change in light levels, and so no adjustment is needed in order for the movement to reach its goal. When a movement is under the guidance of a mechanism, all this entails is that the mechanism stands ready to adjust the movement *if this is required*.

Frankfurt argues that when a movement of the agent's body is guided by the agent, then the agent is moving her own body. We have seen what it is for a bodily

9 (Frankfurt 1978) p. 158

10 Ibid. p. 160

11 Ibid. p. 159

movement to be guided by a mechanism, but what is it for a bodily movement to be guided by an agent*?* Frankfurt's answer is that a movement is under the agent's guidance when it is under the guidance of some mechanism that represents the agent. The activity of such a mechanism 'is, when we are performing an action, our guidance of our behavior.'[12]

On Frankfurt's picture, then, the agent's role in action is not simply to initiate a bodily movement. Rather, the agent has a role throughout the movement's course, monitoring and adjusting the movement if necessary. On the standard picture, the agent's relation to her action is like her relation to a bowling ball she rolls towards some pins: once she has let go of the ball she has nothing left to do but watch it roll towards its target. On Frankfurt's picture, it is as if the agent can run alongside the ball, giving it a nudge if it looks like it is not going to hit the pins.[13]

2.2 The problems with guidance

Frankfurt is right that the standard view of action is unsatisfactory. Agents act over time, and there must be some difference between an action and a mere bodily movement during the whole time the agent is acting, rather than just at the event's initiation. Unfortunately, Frankfurt's account is no closer to providing a plausible picture of the agent's control over her body than the standard view. His account ensures that actions are different from mere bodily movements throughout their unfolding, since they occur under the agent's guidance. But the idea that an agent guides her action gives a picture on which the agent is alienated from the body. We will see that the agent's overseeing and adjusting movements of the body is very different from an agent moving her body. On Frankfurt's picture, the agent's body is

12 (Frankfurt 1978) p. 160. To present a fully worked-out picture of human agency, there are some difficult questions Frankfurt must answer about the nature of this mechanism and, importantly, what justification there is for identifying guidance by this mechanism with guidance by the agent. I will not pursue these issues here.

13 The aim of this example is to provide an analogy to illustrate Frankfurt's notion of the agent's role during her actions: that the agent's role involves monitoring and adjusting a movement rather than simply initiating a movement. Frankfurt does not claim that the agent's guiding her bodily movements is an *action* on the part of the agent, as this would make his account circular. Rather, he claims that a mechanism with which we can identify the agent guides the movement and this somehow constitutes the agent's being in control of her movement throughout her action.

some separate entity whose movements she controls, rather than something that she can move.

We can see the problem with Frankfurt's picture if we consider the similarities between guidance as Frankfurt characterizes it and negative feedback mechanisms. A negative feedback mechanism monitors some variable to ensure that it is within a certain range, and should any deviance be detected the mechanism initiates a means of bringing the variable back within the target range. An example is a thermostat in a home, which monitors temperature in a room to keep it within a certain range and responds to the temperature becoming too high or low by switching the central heating off or on. Importantly, the mechanism's role is simply to bring under control some variable or process separate from itself. The room would still have had a certain temperature had the thermostat not been in operation, and the thermostat's role is simply to regulate the temperature.

On Frankfurt's picture of human agency, the agent's being in control of how her body moves amounts to the agent playing the role of a negative feedback mechanism—she simply regulates the movements of her body. During the time that an action is unfolding, the agent is simply monitoring and adjusting movements of the body to ensure that they take a specific course. The agent is not so much moving her body as overseeing movements that she is not making. This does not seem much like self-movement, being in control of one's body—rather it sees the agent acting on her body as if she were acting on another object.

The problem with Frankfurt's picture is that the agent is alienated from her body—the agent regulates and monitors movements of her body rather than actually moving her body. But it may appear that the problem is not inherent in the notion of guidance itself, and can be solved by adding further conditions to Frankfurt's account. I will show that this is not the case.

2.3 Is guidance really the problem?

I have so far argued that the view that an agent must guide movements of the body gives a picture on which the agent is alienated from the body. Some might think that

this separation between agent and body is not inherent in the notion of guidance itself but only finds its way into the picture because Frankfurt has failed to specify that the agent has a further role *on top* of guidance. We will see, though, that this cannot be right. The problem of the agent's being alienated from her body persists even if we add this extra role of the agent to Frankfurt's account, showing that the alienation problem is inherent in the notion of guidance itself.

The area of the account where extra conditions are seemingly needed is one that is neglected by Frankfurt: the agent's causal production of the movement. We have seen that Frankfurt wants to depart from the standard view that causal history is what makes something an action, but he goes to the extreme in placing no importance at all on how a bodily movement comes about on his account. Frankfurt writes:

> *Given a bodily movement which occurs under a person's guidance, the person is performing an action regardless of what features of his prior causal history account for the fact that this is occurring. He is performing an action even if its occurrence is due to chance.*[14]

This idea certainly adds to the sense that the agent is separate from the body. It allows a picture on which the agent simply finds bodily movements occurring and brings them under control by guiding them towards their end points. So, it may seem that we can avoid the alienation problem by specifying that the agent must not only guide a movement but must also somehow cause it to happen in the first place. But, upon looking at this suggestion, we will find that adding this extra condition does not avoid the alienation problem. The source of the problem is the notion of guidance itself rather than the lack of an additional role for the agent of bringing about the movement.

The first try at an extra condition that requires that the agent cause an action is to require that, as well as guiding the movement, the agent must initiate it. But if we make the agent's initiating movement a condition for action as well as the agent's guiding it, this does not remove the alienation from Frankfurt's view. During the course of the action the agent's role is the same as it was before—the agent is still merely adjusting and monitoring a movement of her body. The fact that the agent

14 (Frankfurt 1978) p. 159

sets the body in motion does not change this.

Another possible condition would require that, in addition to guiding the movement, the agent must be causally sustaining the movement throughout its course. This condition would mean that throughout an action the agent is continuously causally contributing to the continuation of the bodily movement. It requires that if an agent is acting, the continuation of the action depends on the agent's causal contribution. This was not part of Frankfurt's original picture, where the agent's role was to adjust and monitor a movement and not to produce the movement in the first place.

The sustaining condition may initially seem more promising than the initiation condition. Whilst the initiation condition failed because it only concerned the beginning of action, the sustaining condition changes the relation the agent must have to her body *throughout* the action. The agent is sustaining the movement throughout its course, and so it would seem that the agent is more intimately implicated in the movement throughout its course than when she is simply adjusting it. But adding this condition to Frankfurt's account cannot avoid the problems with the notion of guidance. To show why this is, we must first examine how we should understand the idea of the agent causally sustaining a movement, since the notion is not developed enough as it is. When we see what sense this idea could have, we will see that it cannot help with the alienation problem. If an agent is causally sustaining a movement in the sense required to give us a case of action, then guidance can no longer be part of the picture.

What could it be for an agent to causally sustain a movement? One way to think about this is in terms of the movement's depending on the agent's causal role. Without the agent's continued activity the movement could not continue to occur. But this does not provide the notion of sustaining that would be useful in characterising action. To see this, consider an analogy: some trains use a 'dead man's switch' as a safety mechanism—a switch that must be held down by the driver in order for the train to operate. If the driver stops holding down the switch, the engine will be cut and the brakes applied, ensuring that the train cannot keep on traveling if the driver becomes incapacitated. Without the driver's continual pressing down of the switch, the movement of the train could not continue—and so the movement of the train depends on the driver's holding down the switch in a way that fulfils the notion of 'causally sustaining' considered here. But this sense of

'causally sustaining' would not provide a useful characterisation of the agent's sustaining a movement. If an agent's activity is simply allowing a movement to happen in the way that the driver's holding the switch allows the movement of the train to happen, then the agent may still be separate from the body. Instead, what we require is a notion of 'causally sustaining' on which the agent is genuinely continuously *producing* the movement.

However, when we consider what it is for an agent to continuously causally produce the movement, we find that there is no longer a need for the idea of guidance. The agent's continually causally producing the movement must amount to the agent moving and being in control of her body—the very thing we are trying to say more about. It is difficult to see how an agent could be continuously causally producing a movement of her body without this precluding the need for the agent to have some separate role in guiding the movements of her body. So, once we have the idea of the agent producing a movement we no longer need the idea of the agent guiding a movement. Introducing the requirement that the agent causally sustains the movement does not so much solve the problem with guidance as eliminate the need for the notion of guidance altogether.

3.

Steward's Account of Control

3.1 A familiar problem

We have so far seen that Frankfurt's conception of guidance leads to a picture on which the agent is alienated from the body. But since Frankfurt still operates with many of the causal view's assumptions about the nature of action, it is perhaps not surprising that he is led to this problematic conception of guidance. He has a view on which actions are essentially bodily movement events with which agents' mental states interact. His separation of agent and body may be a result of the struggle to accommodate this mental/physical divide.

What *is* surprising is that conceptions of the agent's control very similar to Frankfurt's guidance crop up even in accounts that do not share any of the standard view's starting points.[15] The problematic notion of 'guidance', then, cannot simply be seen as a symptom of the standard view's particular picture of action. In this section I will look in detail at Helen Steward's account of action and show that her attempt to characterise the agent's control leads to the same alienation problem we saw in the previous section.

It is interesting to look at Steward's view because some of her main aims are to minimise the role of the mental in action and to provide a picture of the agent and body as one integrated system—'a whole, functioning animal'.[16] Her starting point, then, is very different from Frankfurt's.

I will explain Steward's view and argue that it does indeed fall subject to the alienation problem. Then, in Section 4, I will draw out a common source of the problem in both Frankfurt and Steward's view.

3.2 Steward's view of actions

A central idea behind Steward's picture of actions is that agency is not a primarily mental phenomenon. Steward argues that proponents of the standard causal view are mistaken to claim that actions are essentially rational or caused by mental states. Rather, the distinguishing feature of actions is that the agent is interjecting into the causal order and making something happen that wasn't determined to happen. Steward argue that acting consists in agents settling what happens. For example, by raising my arm at 10am I settle that my arm will rise at that time. Earlier in the morning, before I acted, it was not determined that my arm would move in that way at 10am—there was no fact about whether or not my arm would rise. I settled that my arm would rise by acting. An action, on this view, is the agent exercising a power

15 Examples of such accounts other than that presented by Steward (2012), are those presented by Di Nucci (2013) and Levy (2013).

16 (Steward 2012) p. 49

to move her body in a way that it did not have to move.[17]

For Steward, the primary way agents settle what happens in the world is through self-movement: "For me an agent is an entity that has a body and can make that body move in various ways and, correlatively, an action is an exercise of this power to make the body (or particular parts of the body) move."[18] Steward looks at self-movement in the stronger sense that we are focusing on in this chapter—not the agent's simply moving by itself but the agent's being in control, or in charge of, its own body.

We might expect Steward's account of what it is for the agent to be in control of her movements to escape the problem of the agent being alienated from the body. She aims to present a picture of agent and body as one integrated system, rather than seeing human agents as beings that can be cleanly divided into mental and physical features: 'the agent is 'a whole, functioning animal whose systems of agent control are various'[19]. Despite this, however, Steward's account ends up being similar in many respects to Frankfurt's, and is subject to the same problem.

We can get an initial glimpse of the similarities between Steward's and Frankfurt's accounts by looking at the passage below. Here, Steward explains that the agent continuously settles what happens throughout the time that she is acting, rather than merely initiating a movement and letting it take its course:

> *My action is the whole, embodied process by means of which the movement is brought about, not just an initiatory push at the beginning of the causal chain, because my role in the relevant causation extends far beyond the role played by the initiatory portion of the chain. This is because I settle things not only by initiating motor activity but also by continuing it; by refraining, for example, from vetoing the original instruction or from altering it in any of the multifarious ways that are constantly open to me. Because these powers of refrainment and alteration are present throughout the whole duration of the action, I am constantly settling what happens from moment to moment, even if I do not in fact exercise those powers of*

[17] Steward uses the term 'power' rather than 'ability', but the powers presented on her account are in many ways similar to the abilities discussed in this thesis. However, Steward argues that the powers agents exercise in acting are 'two-way powers': they are powers to do something or refrain from doing that thing. This is stronger than the idea that exercises of abilities are not necessitated by any circumstances.

18 (Steward 2012) p. 32

19 (Steward 2013) p. 49

refrainment and alteration.[20]

The similarity with Frankfurt's view lies in what it is, according to Steward, for the agent to be continuing a movement. On Steward's picture, what it is for the agent to be continuing a movement is for the agent to possesses the power to stop or alter it. As on Frankfurt's view, the agent here appears to have power over movements of the body rather than the body itself. Like Frankfurt, Steward further claims that the powers need not actually be exercised—what is important is that the agent *could* stop or alter a movement. And also as with Frankfurt's view, we are left with a picture on which the agent is one step removed from her body. The agent's possessing the powers to stop or alter a movement may be enough to ensure that the agent is settling what happens throughout an action, since it is within the agent's power to ensure that any particular movement does not happen. But the agent's possessing these powers is not enough to give us a picture of self-movement on which agent and body are integrated.

Due partly to her move away from seeing the mental as a central to action, Steward is careful to acknowledge the role sub-personal bodily systems play in action. Many bodily movements we make are not consciously controlled by the agent, and Steward even argues that some actions are not intentional and do not involve the agent's conscious awareness at all. The separation of agent and body on her view are cemented in her account of what, if not the agent's conscious involvement, can make these movements the agent's. Steward argues that these movements are the agent's because the agent retains the ability to stop or adjust these movements. In the next two sections I will look more closely at how Steward deals with the idea of sub-personal systems in the body and how this highlights the alienation of agent and body on her view.

3.3 Sub-intentional systems

When agents act, they do not—and cannot—consciously control all aspects of the movements they are making. When I speak I do not consciously control the way my

20 (Steward 2012) p. 46

mouth and tongue move, and when I am typing I do not consider the way that each finger needs to move in order to hit the right keys. These movements are instead taken care of by sub-personal systems in the body, seemingly without any conscious input by the agent. Sub-personal systems, rather than the agent, work out where the fingers need to be to hit the right keys, and how the tongue needs to move to make the right sounds. Given this, Steward worries that we might think that sub-personal systems, rather than agents, control how the body moves. How can I be in charge of how my body moves if the movements of my body are determined by sub-personal systems without my awareness? In order to show that the agent *is* still moving her body in these cases, Steward aims to say 'in virtue of what, if not direct conscious control, these movings really do count as ours'.[21]

Steward's answer is based on the idea that the agent can stop or alter a movement throughout the course of an action. She argues that although the workings of sub-personal systems are not consciously controlled by the agent, sub-personal systems are still 'subordinated' to conscious, personal-level systems. This is because, at any time during an action, the agent can focus her attention on a movement and can alter it or stop it from happening. The exact movements of my fingers when I am typing might usually be due to sub-personal mechanisms, but at any point I can stop or alter the movements of my fingers. For any particular movement of my fingers, then, it is still up to me whether or not that movement happens—since throughout the action I possess the ability to stop or alter these movements. For Steward, these movements are mine, 'not in the sense that I directly and consciously produce the movements from moment to moment, but rather in the sense that I could alter what occurs and in particular could desist from or change the nature of the movements (e.g. from dancing to running) if it became at all important for me to do so'.[22]

Steward further illustrates this view by drawing an analogy between the agent's relation to sub-personal mechanisms and the relation a government minister has to the people working under her:

21 (Steward 2012) p. 50

22 Ibid. p. 51

> *No minister directly controls all the work of a particular department or knows all the details of what is going on within it [...] but if it becomes important, a minister can step in to take a more direct interest in a particular matter. She can more closely monitor the work of the civil servants beneath her, issue directives, insist on changes in working practice, or even, in extremis, take on some of the necessary work herself [...] And so even when she has not been paying attention to some matter or other, we may nevertheless hold her responsible for what has happened in respect of that matter. She could have intervened to change things, even though she did not.*[23]

Although the work is done by civil servants, the minister is in control in virtue of the fact that it is up to her what work the civil servants engage in—she can step in and instruct them to do something else. Similarly, on Steward's view, the agent has the power to change the movements that are being made by sub-personal systems, and this is what makes it the case that the agent is control.[24]

3.4 Sub-intentional actions

In the cases we have considered so far, the agent is intentionally moving her body even though sub-personal systems cause some movements without the agent's awareness: when I type, I am normally aware that I am doing so—even if I am not aware of the exact movements my fingers are making. Steward argues that there exist some actions—'sub-intentional actions'—in the case of which the agent may have no awareness whatsoever of what she is doing. In these cases, personal-level processes are not involved at all. Looking at Steward's account of these cases will reveal a new problem that further shows how the agent is alienated from the body. Steward acknowledges this new problem and offers a suggestion of how it might be solved—but this solution does not eliminate the alienation of agent and body from the account. First of all we must look at what sub-intentional actions are.

During the course of the day I may scratch my head whilst I am thinking, shift the position I am sitting in, or tap my foot along with the music on my headphones—all

23 (Steward 2012) p. 51

24 Steward is not the only philosopher to employ this kind of metaphor in attempting to understand the agent's role in action: David Velleman writes that 'just as the corporate enterprise includes both a basic work activity and the higher-order activity of managing that work… so full-blooded action comprises both a basic activity and the higher-order activity of controlling it.' (Velleman 1992, p. 192)

without any awareness I am doing so. I may only realise I am tapping my foot when someone points out that I am making too much noise. These behaviours do not seem to be goal-directed or done for a reason by the agent—and they certainly do not appear to be intentional. Nevertheless, Steward argues that these are examples of actions. She appeals to the fact that we readily ascribe these behaviours to ourselves as things that we 'do', and we do not lump them in with automatic movements such as breathing or digesting.

Steward's claim is controversial, since it is widely held that all actions must be goal-directed, intentional, and done for reasons.[25] But whether or not Steward is right to claim that there are sub-intentional actions is not important for the purposes of this chapter. What we are concerned with is how Steward views the agent's role in sub-intentional actions, and what it tells us about her picture of the relation between agent and body.

What makes sub-intentional actions different from automatic behaviours? What makes it the case that the agent is moving her body, given that personal level processes are not involved in these movements? As in the case of intentional actions, Steward argues that sub-intentional actions are movements made by the agent because the agent has the ability to stop or alter these movements throughout their course. I may not be aware that I am tapping my foot, but when someone complains about the noise I am making I can bring the movement under my 'direct conscious supervision' and 'stop it in its tracks, reverse, alter, change the direction and speed of, or otherwise affect the motion in question'.[26] These movements are the agent's because the agent is still in charge of whether or not these particular movements occur. There is one main difference between the agent's role in the intentional and sub-intentional cases, on Steward's view: in the case of intentional action, the agent is constantly 'settling' what happens in virtue of her being constantly aware of what she is doing to some extent. In the case of sub-intentional action, the agent has the capacity to bring the action into her awareness and thereby exercise the powers that constitute the agent's being in control of her body: the powers of altering or refraining from altering the movement's course.

25 Brian O'Shaughnessey (1980), who introduced the term 'sub-intentional action', argues that these actions are not immune to all forms of psychological explanation: he claims that they are often causally explained by desires.

26 (Steward 2012) p. 52

The new problem that arises for Steward from her discussion of sub-intentional actions is that, using the tools available on her account, they cannot be differentiated from automatic, non-agential behaviours like breathing and blinking. Christos Douskos points out that most breathing done by agents is automatic and non-agential yet on Steward's view the agent still has the very same powers of alteration and refrainment over her breathing as she does over her sub-intentional actions. Although my breathing may be automatic, 'one can start, stop and alter in multifarious ways one's breathing.'[27] For example, I can hold my breath or breathe more quickly. Even when the breathing is automatic, I still retain these powers over my breathing and so it appears that they should be counted as sub-intentional actions on Steward's view. So, it would seem that to accept Steward's account would be to see our sub-intentional actions as no different from automatic movements.

In a response to Douskos, Steward acknowledges that this is a problem for her account:

> *Breathing, for example, is certainly not normally an agential process; it is an automatic one, which goes on without our needing to guide it, control it or take any agential part in it. Nevertheless, it is possible for me to exert a certain amount of intentional control over my breathing, should I decide I want to do so. I can, for example, hold my breath, at least for a short time, and breathe more deeply, or more quickly or slowly, usually simply by trying to do so. The criterion I suggested, therefore, will not do.*[28]

One way this problem highlights the wider issue with Steward's view is by showing that the condition that the agent is guiding a movement is not sufficient for the agent to be acting. But it is also relevant to the problem of Steward's picture alienating the agent from her body. It shows that Steward's is a strange way to think about actions, since her view forces one to accept that the agent's relation to her bodily movements during her action is the same as the agent's relation to automatic behaviours like breathing and blinking.

The fact that Douskos raises this problem, and that Steward accepts it, also shows how this idea that the agent is separate from her body is taken for granted. There is an easy way that the problem could be avoided, and Douskos' criticism could be

27 (Steward 2013) p. 645

28 (Steward 2013) p. 698

dismissed. But it is not available to Steward because of the way she conceives of the agent's relation to the body.

Douskos' criticism can be rejected because one of his premises is false. It is not the case that 'one can start, stop and alter in multifarious ways one's breathing'. Consider the unconscious, automatic breathing that is happening most of the time when the agent is not paying it any attention. Now imagine that, as this breathing is going on, I decide I will breathe more deeply and I do so. Am I, as Steward and Douskos claim, exercising my power over a token automatic breathing movement—affecting it so as to make it a movement of breathing more deeply? Or, alternatively, am I just performing a different type of movement altogether—one that is an action? My contention would be that it is the latter. I do not have the power to 'start, stop and alter' a movement that is automatic, unconscious breathing. I simply have the power to start breathing where this breathing is an exercise of an ability, which means that an action happens rather than the automatic movement that would otherwise have happened.

So why does Steward not deny Douskos' claim that I can 'stop, start or alter' a non-agential automatic token movement of breathing, instead of accepting his criticism? It is seemingly because her view of the agent's relation to her body means that she *does* think that an agent who decides to breathe more deeply is altering an already-occurring movement rather than simply bringing about another type of movement. If one thinks that an agent's acting involves the agent's affecting (or having the opportunity to affect) the movements of her body through focusing conscious attention on them, then one will think that when one breathes more deeply or quickly, one is affecting an already occurring movement rather than simply moving one's body in a certain way. The fact that Douskos accepts this kind of picture as well, as shown by his criticism, illustrates how easy it is to slip into this way of thinking about the agent's relation to the body.

In response to Douskos' problem, Steward modifies her view. It is important to look at whether or not this modification eliminates the alienation problem. Steward sees the modal nature of her account of sub-intentional actions as what gives rise to Douskos' criticism. She suggests that it is not that the movement *could* be brought into her awareness and thereby under the agent's control that makes it an action—as her account previously stated—but the fact that it *is being controlled already* in much the

same way as an intentional movement.[29]

A preliminary suggestion by Steward is that we find a minimal kind of agential awareness in sub-intentional actions that does not amount to full conscious control but still allows us to say that sub-intentional actions are somehow 'meant' by the agent: 'One may not intend to be fiddling with one's jewellery. But nevertheless, there is a sense in which the movements, one might think, are *meant*.'[30] In this case, Steward need not rely on the idea that these movements *could* be brought into the agent's awareness. She can require that in a sub-intentional action this 'meantness' is present, and that the agent's awareness must be somehow *actually* involved in the production of the movement. The agent's awareness—this 'meantness'—is present in cases of sub-intentional action, but not present in the cases of automatic movements like breathing.

The idea that there are degrees of awareness could perhaps be made plausible, and Steward's addition of this condition does help her avoid Douskos' criticism: it allows her to differentiate between sub-intentional actions and automatic behaviours, and so exclude automatic behaviours like breathing from the account. The problem of alienation, however, persists. Steward manages to exclude automatic behaviours from being counted as actions by bringing her account of the agent's involvement in sub-intentional actions more into line with her account of intentional actions. But we have seen that the alienation problem is present in Steward's picture of intentional actions too, since on her view the agent's actually controlling her action amounts to the agent's having the power to oversee and adjust movements of the body rather than genuinely moving the body. What it is for the agent to be consciously controlling an action is for her to have 'powers of refrainment and alteration' over a movement, rather than for the agent to be directly moving her body. Making the sub-intentional case more like the intentional case, then, cannot avoid the alienation problem.

29 (Steward 2013) pp. 698-699

30 Ibid. p. 699

4.

Control and Ability

4.1 How should we characterise control?

Both Frankfurt and Steward attempt to give accounts that avoid a separation between agent and body during action, but their characterisation of the agent's being in control of her body as a matter of monitoring and causally affecting movements of the body leaves the agent more alienated from the body than ever. They run into problems because both are drawn to the thought that an agent's being in charge of her body must involve possessing a power over movements that are produced or sustained by mechanisms separate to the agent. But we can avoid this problem if we abandon the conception of the agent's relation to the body that underlies their accounts. Once we stop seeing the agent's body and the physiological systems involved in the movements that the agent makes as essentially external to the agent, this notion of the agent having a power over something separate is no longer needed.

For someone like Frankfurt, who takes as his starting point the reductionist idea that actions are identical with non-agential bodily movement events, abandoning the problematic assumption will be easier said than done. The reductionist approach itself is one that tends towards a picture on which the agent is separated from the body. The agent is seen, first and foremost, as a primarily mental and rational being and the movements of the body as simply physical goings-on. The reductionist's aim is to answer the challenge of how the former must causally affect the latter in order for there to be an action. When the pieces of the puzzle are set up in this way, we are already left with a picture on which the agent and body are essentially separate. We cannot throw out this assumption without abandoning the whole reductive approach.

Steward, on the other hand, is someone who has already rejected the reductive approach and holds the view that actions are essentially agential exercises of power by the agent—but she still has trouble fully abandoning the view of agent as

essentially separate from the physical body. What shows that Steward has not fully abandoned this view is her worry that much that goes on in the body during actions is the work of sub-personal systems and is not accessible to the agent's conscious awareness. In these cases, Steward believes we need to ask: 'in virtue of what, if not in virtue of direct conscious control, do these moving really count as ours?'[31] My contention is that, once one has abandoned the reductive approach, there is no reason to see the physiological systems at work in the body as the workings of something other. These movements are mine, and I am in control of my body, in that I am moving my body and these movements are part of my exercising my abilities. The workings of sub-personal mechanisms are not separate processes in which I need to intervene—they are simply part of my moving my body. I bring about the workings of the sub-personal processes by moving my body, insofar as they are necessary for my body to move in the way that I am moving it.

Steward writes that an agent continues making a movement of her body by 'refraining from vetoing the original instruction'.[32] But this gives the impression that when we act, we issue an instruction to some external power and let it carry out the process. On the view that I am proposing, the agent does not continue a movement by refraining from vetoing an instruction—she just continues a movement by continuing to exercise her ability to move her body. Similarly, if I am moving my arm and I want to stop moving it, I do not need to veto an instruction—I just need to stop moving my arm. If I want to move my arm in a different way to the way I am currently moving it, I need not intervene in or adjust some separate process that is currently underway—I just need to move my arm in a different way. The way in which I control how my body moves is not by initiating, overseeing and intervening in the movements that my body makes. I control my body simply by exercising my abilities.

It might be thought that my conception of the agent's relation to her body fails to recognize the points that Steward is trying to address. What Steward seems to be concerned with in her account is the thought that in some cases, when sub-intentional systems are at work without any conscious input from the agent, our bodies do appear to be 'other'—and we cannot just say that these movements are

31 (Steward 2012) p. 50

32 Ibid. p. 46

mine in virtue of the fact that *I* am moving my body. It is in these cases that we need to ask: 'in virtue of what, if not in virtue of direct conscious control, do these moving really count as ours?'[33]

My claim, though, is that it *is* an acceptable answer to Steward's question to say that these movings are not 'other' and are ours in virtue of the fact that they are part of our exercising the ability to move our bodies. There is little more to say about how the agent controls these movements and how she relates to the workings of these mechanisms apart from what has already been said: that they are the agent's because the agent is exercising her ability to move her body.

Unless Steward is asking the question asked by many reductionists—how we can specify, in non-circular terms, the difference between the mere movements of our bodies and those movements that are actions—there should be no issue with answering her question in the way that I have. We have seen that those who take the standard approach are wrong to start off with a bodily movement event that could be an action or a 'mere bodily movement', and then ask what kind of causal relation the agent must have to it in order for it to be an action. It is equally wrong to start off with the activity of an agent's sub-personal mechanisms, and ask what in what supervisory relation the agent must stand to this activity in order for it to count as the agent's.

Acknowledging the existence of sub-intentional actions is important, because they demonstrate interesting features of agency. They show that the agent's exercising a power to move her body need not always be conscious, intentional or rational. But, just as in other cases of action, what makes the movements involved *mine* rather than the work of some external power is that *I* am moving my body. We should not try to say more where there is nothing more to be said.

It is worth noting that, in an earlier paper on sub-intentional actions, Steward takes a similar line to this one: she writes that 'it is better to say a small amount that is true, then a very great deal that is false.'[34] In this paper, she avoids the kinds of metaphors that give a picture on which the agent is separate from her body, and claims only that sub-intentional actions are actions because they involve exercises of a power by

33 (Steward 2012) p. 50

34 (Steward 2009) p. 310

the agent. In elaborating on this in the later *A Metaphysics for Freedom*, however, Steward is drawn into asking the further question of why the movements brought about by sub-personal mechanisms are *ours.*[35]

In allowing that this question can be asked even once we have accepted that actions are exercises of an agent's power (or ability), Steward is led to a picture on which the workings of sub-personal mechanisms are 'external' and separate from the agent. She is led to drawing a distinction: between movements that are under our direct conscious control and are ours simply because they involve our exercising a power to move our bodies; and movements that are not under our conscious control but are brought about by sub-personal processes, that are ours in virtue of the fact that we can oversee and intervene in them.

Conclusion

An agent's self-movement—moving her own body and being in charge of how her body moves—is essential to her exercising abilities. But what we have now seen is that attempting to explicate self-movement without reference to the notion of an agent's exercising an ability will be unsuccessful. We have looked at two related questions about self-movement, and in both cases we have seen that the answers are less informative than those asking the questions might have hoped.

The first, general, question—asked by both Frankfurt and Steward—was: what is it for the agent to be in control of the movements of her body? The answer to this question is simply that the agent is in control of her movements when she is exercising her abilities. The notions of self-movement and exercise of ability are too closely connected to say anything further that does not separate the agent from the body.

This first question may have a less informative answer than was expected, but there is nothing mistaken about posing the question itself. This is not the case for the more specific version of this question raised by Steward: why should we think that the agent, rather than sub-personal mechanisms in the agents body, is in charge of

35 (Steward 2012)

the agent's movements when she is acting?

We saw that it only makes sense to ask this second question if one's starting point is the assumption that the bodily mechanisms at work when the agent acts are not essentially part of the agent's exercising an ability. We must then, as Steward does, come up with a picture on which the agent inserts herself into the activity of these sub-personal mechanisms by overseeing and guiding their activity. The answer is to reject the picture of agency that inspires the question. Our starting point must be a picture of the agent as an embodied being, the activity of whose bodily mechanisms when she acts is simply part of her exercising her ability rather than a separate process she must insert herself into.

Chapter 5

Self-movement and Other-movement

In a recent paper, Anton Ford has issued a challenge to the idea that self-movement is central to human agency—a challenge that I will respond to in the present chapter. We saw in Chapter 4 that it is difficult to say much about what it is for agents to be in charge, or in control, of their bodies that is both true and informative. The idea of an agent's being in control of her body is too closely tied to the idea that human agents exercise abilities to move their bodies. Responding to Ford's challenge will open up a way to say more about the role of self-movement in human agency, and will allow us to see how the account of dispositions developed in Chapter 3 can be brought to bear on an account of human agency.

Ford argues that a preoccupation with self-movement leads many philosophers towards a distorted view of the way human agents make their way in the world. An exclusive focus on self-movement leads them to miss the fact that human agents typically exercise powers over objects other than their own bodies. I will suggest that many of Ford's concerns about an undue emphasis on self-movement in theories of action are legitimate. But I will argue that Ford's neglect of self-movement in his own account leads to a picture of human agency which contains a gap: Ford leaves us with no way to articulate the embodied nature of human agency. It is through developing a view of self-movement that fills the gap left by Ford, that we will arrive at a better understanding of the role self-movement plays in human agency, and make the sort of progress that seemed to be ruled out when we concentrated on the notion of bodily control.

I will first set out Ford's view, and indicate the aspects of his account that we must accept. I will argue that we can accept these claims without denying that self-movement has a central role in human agency, before showing exactly what exactly is missing from Ford's account through his failing to pay due heed to the notion of self-movement. Filling this gap will involve applying the account of dispositions developed in Chapter 3 to the case of abilities.

1.

Corporealism and Materialism

1.1 Corporealism

Ford sets out to challenge what he takes to be a consensus in the philosophy of action: the idea that human agency is fundamentally a matter of self-movement. When philosophers set out to investigate human agency, they investigate self-movement—agents' moving their bodies.[1] And they take it as given that in providing an account of self-movement they are providing an account of human action. Ford argues that, despite its going virtually unquestioned, this approach leads to philosophers presenting a distorted picture of the nature of human agency.

The view that Ford criticises is inherent in Donald Davidson's famous claim that 'our primitive actions, the ones we do not do by doing something else, mere movements of the body—these are all the actions there are. We never do more than move our bodies: the rest is up to nature.'[2] The thought expressed by Davidson here is that when human agents do things like switching on lights, opening doors or kicking footballs, their contribution to what is happening their moving their bodies. Everything else that happens is then left 'up to nature' in the sense that it is simply a causal effect of agents' moving their bodies. Of course, Davidson does not go so far as to claim that statements like 'she kicked the football' or 'he opened the door' are false. He maintains that the bodily movement events that are our actions can be redescribed in light of their effects: my movement of my leg can be described as my kicking of the football because it brings about the football's being kicked. Still, on Davidson's view, my action is a matter of my moving my body.

The sentiment of Davidson's claim that 'we never do more than move our bodies' is shared by many others—even those who reject Davidson's causal theory of action.

---

1 Not all philosophers who focus exclusively on bodily movement would describe an agent's moving her body as self-movement. Those who are apt to think in dualistic terms may conceive of the body as something 'other'. Moving one's body, on this view, would be a matter of moving something other rather than moving oneself.

2 (Davidson 1980) p. 59

Adrian Haddock rejects Davidson's view in favour of a 'disjunctivist' theory of action, the details of which I will not go into here.[3] What is significant is Haddock's acknowledgement that the one thing he and Davidson can agree on is that our starting point in investigating human agency must be the thought that actions are bodily movements.[4] Helen Steward, who takes it that actions are exercises of powers by the agent, also privileges the agent's moving the body. She argues that 'an agent is an entity that has a body and can make that body move in various ways and, correlatively, an action is an exercise of this power to make the body (or particular parts of the body) move.'[5] Ford terms the approach of these philosophers 'Corporealism'. According to Corporealism, agency is a power to move one's body.

Ford gives the Corporealist credit for moving away from an even more problematic picture of human agency: Volitionalism. The Volitionalist sees human agency not as a power to move one's body, but as a power to perform mental acts of willing, termed 'volitions'.[6] For the Volitionalist, human agents' fundamental contribution to the things they do is their performing mental actions, which then cause bodily movements and further effects in the world. The Volitionalist version of Davidson's famous claim would be: 'We never do more than perform mental acts: the rest is up to nature.' For the Volitionalist, an account of human agency is an account of human agents' mental acts and the way in which these bring about further effects.

Ford applauds the Corporealist for avoiding the Volitionalist picture. But he argues that the Corporealist approach is problematic in similar ways to the Volitionalist approach. Volitionalism implausibly restricts human agency to the mental, separating human agents from their bodies and the rest of the world, which on the Volitionalist view they can only affect indirectly. Similarly, Corporealism implausibly restricts human agency to the body, separating human agents from the objects they interact with every day, which on the Corporealist view they can only affect indirectly. Ford argues that we should see human agency not as a matter of mental acts or of bodily

3 (Haddock 2005)

4 Ibid. p. 162: 'this thought is the central insight of the so-called 'standard causal story of action', a story with which my account has little else in common.' Quoted in (Ford 2016) p. 21

5 (Steward 2012) p. 32

6 Examples of the Volitionalist view can be found in (McCann 1974), (Smith 1988) and (Pietroski 1998).

movement, but as a matter of human agents directly interacting with things in the world.

1.2 Materialism

Ford contrasts Corporealism, on which agency is a power to move one's body, with a view he terms Materialism, according to which 'agency is a power to change, neither one's mind, nor one's body, but whatever it is—whatever material—one changes intentionally."[7]

The Materialist avoids separating human agents from the world they find themselves in. On this view, human agents' power over the objects they interact with is not mediated by their moving their bodies: 'the immediate object of agency is precisely *not* oneself: it is neither one's body, nor one's soul, but an extra-corporeal object or a second bearer of the power'.[8] When I kick a football, it is not the case that I move my foot, and my foot's moving then causes the football to move. Rather, when I kick a football, I directly move the football. The Materialist maintains that an account of self-movement misses out much that is central to human agency, and argues that we must understand human agency as a matter of humans interacting with things other than themselves.

The Materialist position throws out Davidson's slogan 'We never do more than move our bodies', and adopts a different one, from Elizabeth Anscombe: 'I do what happens.'[9] Ford argues that the transactions with other objects that are crucial to human agents' making their way in the world—pushing, pulling, lifting, drinking—cannot be split up into what the agent does and the effect this brings about in the object she is acting on. When a human agent takes part in a transaction, what she does and the change this brings about in the object she is acting on are identical. So, when I open a door, my opening the door is identical to the door's opening—they are one and the same process. The idea behind 'I do what happens' is that what I am

7 (Ford 2016) p. 4

8 (Ford 2013) p. 601

9 (Anscombe 1957) p. 52

doing when I transact with an object is identical with the process of change that is being brought about.

I agree with the view of transaction that Ford presents here—it is true that we cannot split up an agents' transaction with an object into two parts. Instead we must see the agent's acting on the object and the object's being acted on as one and the same process. To see why one should accept this view of transaction, even if one does not initially share Ford's aversion to Corporealism, we can look at an argument presented by Jennifer Hornsby.[10]

Hornsby shows that it is implausible to split actions up into what the agent does and what happens as a result of the agent's action.[11] The reason for this is that, in many cases, what an agent is doing and what is happening to the object she is acting on are simultaneous: when my action is under way, there can be no separate event that my action is bringing about. Let us consider an example. On a view which splits up transactions into two parts, my action of turning on the tap consists of an event of my moving my body, which brings about an event of the tap moving. But this cannot be right: as I am turning on the tap, there is no subsequent event of the tap moving that my bodily movement is causing. When I turn on a tap, the tap's moving is simultaneous with my moving my hand—both are ongoing at the same time. My moving my hand and the tap's moving are part of the same process—not two separate events linked by causality.[12]

So, we should accept that human agents 'do what happens': they transact with the objects they act on, rather than merely causing changes in objects by moving their bodies. But this does not mean we must accept Ford's Materialist view unreservedly: we must be careful at this point not to conflate two distinct claims that make up the view Ford presents. One claim is that agents 'do what happens'. The other, less straightforward claim, is that philosophy of action unduly focuses on self-

10 (Hornsby 2011)

11 Ibid. pp. 107-108

12 There is a respect in which Hornsby's view of transactions differs from Ford's. Hornsby accepts that there are some cases in which my interactions with other objects are mediated: whilst if I pull a door open the door's opening is identical to my opening it, if I press a button to open the door there is no transaction between me and the door (Hornsby 2011, p. 122). Ford, however, holds the view that even in the latter case I am directly transacting with the door (Ford 2016). This difference between the two views is not important to the present discussion.

movement—that it is mistaken to take self-movement to be the fundamental form of human agency. Ford does not go far enough towards emphasising that these two claims are independent of one another. I will argue that we should accept the former, but that we should not be so quick to accept the latter.

Of course, if we accept that agents 'do what happens' this lends weight, in various ways, to the claim that more emphasis must be placed on other-movement than on self-movement in theories of human agency. It shows that the body is not the only thing over which the agent has direct power, and involves a rejection of the picture of causation that leads those like Davidson to claim that 'we never do more than move our bodies'. But, as we will see in Section 2.3, even when we accept that agents 'do what happens' there are other reasons to think that self-movement deserves its prominent place in theories of human agency.

The first step, however, will be to show that the two claims that make up Ford's Materialist view really do come apart. I will show that even when it is accepted that agents 'do what happens', Ford's arguments fail to show that an emphasis on self-movement is mistaken. I will then look at what exactly is missing from Ford's Materialist view, before going on to suggest how the gap Ford leaves can be filled.

2.

What Is Missing From Ford's Account

2.1 Arguments for the focus on transaction

Ford presents two main arguments to support his claim that the Corporealist position should be rejected. His main target is the theory of human agency presented by Helen Steward in *A Metaphysics for Freedom*: Ford argues that Steward's theory

wrongly focuses on self-movement.[13] But the arguments Ford uses to criticise Steward's view can apply more generally to any Corporealist theory of action.

The aim in looking at Ford's arguments here is not to decide whether or not we should accept the existence of transaction and the idea that agents 'do what happens'—this we have already accepted. The aim is, rather, to see whether the case Ford makes would prevent us from accepting both that agents 'do what happens' and that an emphasis on self-movement is required in a theory of human agency. Ford succeeds in showing that philosophers of action must acknowledge and investigate transaction if they hope to present a theory that truly reflects the nature of human agency. But the arguments do not rule out the need for an emphasis on self-movement as fundamental to human agency.

Ford's first argument is that 'Corporealism does not represent the point of view of common sense'.[14] His thought is that in most cases where we interact with other objects, we think of what we are doing in terms of how we are moving the object rather than how we are moving our bodies.

In support of this idea, Ford elsewhere argues that even Davidson himself is aware that his claim, 'we never do more than move our bodies', is counterintuitive: Davidson acknowledges that this claim will be met with a 'shock of surprise'.[15] Ford takes this to be an admission that the view on which agency is merely a matter of bodily movement fails to reflect the way we think about our interactions with the world.[16]

Davidson's admission is telling against the standard event-causal account, which denies the existence of transaction altogether, but not against the view that I am proposing. The 'surprise' is at the claim that we never transact at all—a claim that we should reject. It need not be surprising that although we do transact, and often think of what we do in terms of transaction, self-movement must still play a prominent role in an account of action.

13 (Steward 2012)

14 (Ford 2013) p. 609

15 (Davidson 1980) p. 59

16 (Ford 2016) p. 3

Ford's second argument is based around a thought experiment.[17] He aims to show that we lose nothing in privileging other-movement over self-movement, by showing that a view that focuses on self-movement gives a counterintuitive picture of the nature of our agency. Ford asks us to imagine a person floating through space or suspended in jelly, who has no opportunity to interact with any other objects but can still move her own body: she can still raise her arm or scratch her head. Ford argues that if human agency is a power of self-movement, then nothing essential to agency should be missing from this case—but he goes on to claim that something essential does appear to be missing, and that this demonstrates that self-movement is not the 'fundamental form' of human agency.

Ford claims that something essential is missing in this case because 'the movement proper to a human body, and therefore to a human being, essentially involves extra-corporeal objects.'[18] This, he argues, is because our body parts, such as our limbs, have natural functions—and the functions of our limbs, and of our hands and feet, are to interact with other objects: 'The human body is for a world distinct from the human body, and human bodily movement is for transacting with that world.'[19] In the case where a human agent is floating in space with no opportunity to act, something essential to their agency is missing.

It can be acknowledged that there is something missing in some sense in Ford's thought experiment: the agent floating in space is unable to act in many of the ways that we normally do act. But Ford wants a stronger conclusion. The fact that the agent is unable to act in many of the ways that we normally do is supposed to show that something essential to agency itself has gone missing.

Ford's argument here relies heavily on intuitions that those opposing him do not share, so that it is unlikely to be persuasive to anyone who does not already hold his view. For example, Steward contends that there is nothing essential to agency lacking in the situation presented in the thought experiment, despite the fact that the agent has the opportunity to move only her body.[20] Ford's claim is that the human

17 (Ford 2013) pp. 609-610

18 (Ford 2016) p. 13

19 Ibid.

20 (Steward 2013) p. 689

body's function of interacting with worldly objects ensures that something is missing in his example. But the claim is simply a statement of his view rather than an argument. It is unclear why the human body's having the function it does must entail the conclusion that something essential to agency is missing in the case presented.

Nothing in Ford's arguments has shown that accepting that agents 'do what happens' requires prioritising other-movement over self-movement in theories of action. Even when we accept that agents transact, this leaves it open that an account on which self-movement is seen as fundamental to human agency can provide a plausible picture of human agency. We must now turn to the positive reasons for supposing that, although agents 'do what happens', an account which has more to say about self-movement is required. We will look at what is missing from the Materialist view due to its failure to acknowledge the central role of self-movement in human agency. Before explaining the genuine problem with the picture Ford presents, I will put to one side a mistaken, but initially appealing, argument that self-movement is fundamental to human agency.

2.2 An unsuccessful argument for focusing on self-movement

Even if one accepts that agents 'do what happens', there is still an element of the idea that 'we never do more than move our bodies: the rest is up to nature' that may seem appealing. Clearly it is false that we 'never do more than move our bodies'—we have seen that agents often do more than move their bodies, in that they transact with other objects. But it may appear that there is a weaker version of this claim that we should accept. It appears we have more control over our bodies than we do over the other substances we move. The thought here is that we can ensure our bodies move in certain ways whilst our attempts to move other substances require the 'cooperation' of these substances. It is up to me whether I raise my arm, but it is not up to me whether I open a door—since the door may be locked. Our interactions with other objects seem to leave more 'up to nature' than do movings of our bodies.[21]

21 An argument along these lines is made by Frederick Stoutland (2011b, p. 313).

The thought that it is only when we are moving our bodies that we are fully in control may suggest that if we want to give an account of what it is to be an agent, we should focus on agents' powers of self-movement. But this argument is ultimately unsuccessful. There are two objections that demonstrate this.

Firstly, it is mistaken to think that our bodies' moving in the way that we want is entirely up to us. Just as something can 'go wrong' to prevent an agent from opening a door—such as the door's being locked or the handle's being too slippery—something can 'go wrong' to prevent an agent moving her body—such as her muscles not working properly or her bodily parts being paralysed.[22] There is much that can go wrong within our bodies that would prevent us being able to move them in the way that we intend. There is therefore a strong similarity between the level of control we have over our moving our bodies and the level of control we have over our moving other objects.

Even so, one might try to preserve the thought that we have more control over self-movement, by arguing that self-movement is still *more* reliable than other-movement. Perhaps we can be more confident that our bodies will move in the way that we want than that other objects will move in the way that we want. It might be suggested that cases where something goes wrong with my moving my body are very rare compared to cases where door handles are broken or light switches do not work. But even this weaker claim does not hold up.

A key aspect of our everyday practical thought is the assumption that things like doors and light switches will behave as we expect them to—we could not get anything done if we were in a constant state of doubt about the way in which other objects will behave.[23] In our everyday thought, we think of our bodies and other objects as being equally reliable—in that we simply assume that things will behave in the way we expect them to, just as we simply assume that our bodies will function properly when we move them.

22 This point is made by Will Small [forthcoming].

23 The discussion in Chapter 3, Section 1.2 touched on this issue. There, I argued that we can know that the objects we interact with will behave in certain ways, despite our awareness that they are not necessitated to act in these ways in certain circumstances.

What this first of the two objections shows, then, is that we generally think of ourselves as being able to ensure that the objects we interact with move to the same extent that we are able to ensure that our bodies move. In neither case is success a given. But we generally see both self-movement and other-movement as being equally 'up to us', rather than going through life in a constant state of doubt about whether we can actually do the things we intend to do.

The second objection concerns a problem that is due to the argument being rooted in the Davidsonian standard story of action. The standard story, with its separation of bodily movement events and the movements of other substances that are caused by them, lends itself to the thought that I may ensure that the former event happens but not that the latter event happens. The idea that there is a one-way relationship of causation between my bodily movement and movements on the part of the object I am acting on rules out that anything could impede or interfere with my moving my body. But once we take the view that we transact with other objects, we see that our bodily movements and the movements of the substances we transact with are much more intimately connected. It then becomes clear that there is no significant sense in which it is more 'up to me' that my body moves than that the objects I am acting on move.

Imagine that I try to lift a heavy rock. We have seen from Hornsby's argument in Section 1.2 that when I lift the rock, we do not have one event (a movement of my arm) and subsequently another event (a movement of the rock). The mistaken thought that we have two events happening one after the other takes us to the idea that when I try to lift a heavy rock, I first move my body in a certain way and then, if I am successful, the bodily movement will cause a movement of the rock. It suggests that if I fail to lift the rock because it is too heavy, what has happened is that I have moved my body in the way that I intended, but this failed to cause the subsequent event of the rock moving.

But in fact, as we saw from Hornsby's argument, if I move the rock, the rock's moving and my moving it are happening simultaneously. When I try to move a heavy rock, I do not move my body in the way that I would move it if I were lifting it, and then see whether the rock moves. If the rock is too heavy for me to move, I am unable to move my body in the way I would move it if I were lifting it. In many cases, we require the cooperation of the objects we are acting on in order to move

our bodies in the first place. This gives us another reason why it is mistaken to think we can ensure that we move our bodies whilst we cannot ensure that other things move in the way that we want when we act on them.

These two objections show that, despite its initial appeal, the thought that we can ensure we move our bodies is not a good reason for a greater focus on self-movement. We must look elsewhere to see where the gap in Ford's Materialist account truly lies.

2.3 What Ford's account is missing

The true reason Ford's Materialist account needs modifying to include more of a focus on agents' moving their bodies is that his account leaves us with no way to articulate the distinctive embodied nature of human agency. We are biological beings and for the most part—even in cases where we transact with other objects—we exercise our agency through our bodies. In attempting to move away from the focus on self-movement, Ford's view leaves us with no satisfactory way to express the role self-movement undeniably plays in our agency.

Will Small makes a similar claim against Ford.[24] However, Small's diagnosis of the problem with Ford's view misses the mark to a certain extent. Looking at why this is will assist in clarifying my own criticism of Ford.

Small begins his criticism of Ford's view by asking us to compare one person's raising her arm and another person's raising her hat. Small argues that there is a clear difference between the two actions: 'no one can "just" raise a hat as he can (assuming he is not paralyzed etc.) "just" raise his arm'.[25] But, Small argues, on Ford's account we cannot express this idea. On Ford's view, when I raise a hat I am doing so non-mediately. This leaves us with no way to express the idea that when I move my body I 'just' move my body in a way that I cannot 'just' move my hat. As

24 (Small [forthcoming])

25 Ibid. p. 17

Small puts it: 'Presumably one cannot do anything less mediately than by doing it non-mediately.'[26]

For Small, the problem with Ford's Materialist view is that it does not account for the difference in the way transactions and self-movement are mediated. With his example of raising one's arm and raising one's hat, Small aims to highlight the intuition that our transactions with objects are still in some way more 'mediated' than our self-movements. He then claims that Ford cannot account for this: Ford's view does not allow us to say in what sense our self-movements are more direct than our interactions with other objects.

The problem with Small's criticism of Ford is that it makes it look as if Ford should simply reject the demand that it imposes on him. Ford can simply accept that his view does not acknowledge that self-movement is more direct than transaction, since the conclusion he is arguing for is that there is no difference in how 'direct' these two actions are. Ford's aim is to show that a human agent's acting on an object involves a direct and unmediated exercise of power. So he does not need to explain 'how one cannot do anything less mediately than by doing it non-mediately': his claim is that both raising one's arm and raising one's hat are equally unmediated.

The only thing that could salvage Small's criticism is our finding that we have a strong intuition that there is a difference in directness between raising one's arm and raising one's hat. This strong intuition would require explanation by Ford. However, it does not seem that the intuition Small appeals to is strong enough for it to present a problem for Ford's view. Perhaps there is something appealing about the thought that I do not 'just' move my hat in the way that I 'just' move my arm. But we saw in Chapter 4 that we should be wary of giving too much weight to these vague intuitions. There, we saw that the idea that we are the 'overseers' of our bodily movements is easy to slip into, despite its implausibly separating the agent from the body. Similarly, the idea that self-movement is more direct than other movement appears easy to slip into although it implausibly separates the human agent from the world she interacts with.

26 (Small [forthcoming]) p. 17

My claim against Ford's view is not that it cannot explain the way in which self-movement is more direct than other-movement. My claim is that Ford's picture does not allow us to articulate the embodied nature of human agency. With Ford's Materialist picture we lose the sense that human agents are able to do what they do because they are beings with powers of self-movement. We lose the sense that self-movement has a role in everything that we do, even in transaction.

Ford has presented a view on which, in addition to engaging in self-movement, human agents engage in other-movement. He sees these as two different forms of action: 'moving oneself from place to place, or changing (simply) the position of one's limbs, is fundamentally different from transacting with something else.'[27] But what we lack from his picture is an acknowledgement that self-movement plays a role in our transactions with other objects, and that this need not contradict the thought that 'I do what happens'.

My criticism of Ford's view, unlike Small's, does not demand a change in the substance of what Ford says about transaction. Rather, it simply highlights that there is something more that needs to be said about self-movement on Ford's picture. In the next part of this chapter I will apply the multi-track picture of dispositions developed in Chapter 3 to the case of abilities, and slot this into Ford's view of transactions. This allows us to see how we should articulate the role that self-movement and abilities play in all of the transactions human agents take part in.

3.

The Role of Self-movement

3.1 Modifying the Materialist position

I will argue that the central role of self-movement in human agency can be properly accommodated in the Materialist picture when we accept that all transactions are

27 (Ford 2014) p. 24

exercises of bodily abilities. When agents open doors, fry eggs, or prune roses, the abilities they exercise are abilities to move their bodies—the same abilities that are exercised in self-movement are those that are exercised in other-movement. It is in this sense that self-movement plays an important role in human agents' transactions.

It is initially difficult to see how it could be the case that transactions are exercises of bodily abilities, without this requiring a return to thinking of our interactions with other objects as being split into two parts: self-movement by the agent and the effects that this self-movement brings about. Looking more closely at questions about the individuation of abilities will show how we can accept that transactions are exercises of bodily abilities without weakening Ford's view that agents 'do what happens'.

In his discussions of self-movement and other-movement Ford does not restrict himself to a particular view of the nature of causal powers and their manifestations. He takes it, rightly, that his view is compatible with several different approaches to the causal powers of human agents and the individuation of these powers. We have, however, developed an account of dispositions and their manifestations that deals with questions of individuation—which we can now apply to the case of abilities.

In Chapter 3 we adopted a multi-track view of dispositions, and given the other similarities between dispositions and abilities it is natural that this multi-track view should be applied to abilities as well. There are two main aspects of this view that will help us to see how all of our transactions are exercises of bodily abilities. The first is the idea that an ability can have more than one exercise type, and the second is a principle for grouping these exercise types.

3.2 Mutual manifestations and exercise types

Let us review the conclusions drawn about dispositions in Chapter 3. We arrived at a picture on which objects mutually manifest their dispositions with the dispositions of other objects. Dispositions manifest with their manifestation partners in a single process in which both objects take part, and the same disposition may produce different effects with different manifestation partners. We saw that it follows from

this view that dispositions must be 'multi-track'—each disposition may have many different manifestation types, and how a disposition manifests depends on the nature of the other dispositions it is manifesting with.[28]

Having adopted such a picture of dispositions, we should also adopt a view of abilities on which they mutually manifest with other powers: there is little reason to think that abilities are so different from dispositions that they do not also mutually manifest with other powers in a single process. The one key difference in the case of abilities is that being in the presence of their manifestation partners does not add to the likelihood of an ability being exercised: the fact that I am in a situation that gives me the opportunity to raise my arm does not tell us anything about whether I will in fact do so. But this does not affect the fact that, like dispositions, when abilities are exercised their exercise depends on other objects manifesting their powers.[29] For example, the process of my stretching an elastic band is an exercise of my ability but also a manifestation of the elastic band's disposition of elasticity. If the elastic band lacked this disposition, I would be unable to exercise my ability.

It is worth noting that a picture on which abilities mutually manifest with other powers reinforces the idea that when human agents act they are transacting with other objects. It shows that our bringing about changes in the world is not 'one-way', with human agents acting and their actions causing effects in objects. Rather, an agent's acting and an object's being acted on is a single process that depends on both the human agent and the object she is acting on possessing the necessary powers.

With this picture in mind, we can now make the first step towards seeing how it can be the case that all of human agents' transactions are exercises of bodily abilities.

One factor that leads to its initially appearing implausible that transactions are exercises of bodily abilities is our 'pre-theoretical' view of abilities and their exercises, which portrays abilities as single-track powers. It seems natural to speak as if, for

28 We were led to this multi-track picture of dispositions by the idea that the manifestations of dispositions are the effects they have a part in producing, rather than their contributions to the effect. Since each disposition produces different effects with different manifestation partners, the view that manifestations are effects entails that each disposition has more than one manifestation type.

29 I will focus here mainly on examples of human agents interacting with objects. These examples involve a human agent's ability and an object's disposition jointly manifesting/being exercised in one process. When two human agents interact with one another it will be the case that there are two abilities mutually exercising in a single process.

every action type, there is a corresponding ability of which actions of that type are exercises. An agent's speaking French is an exercise of her ability to speak French; an agent's making a cup of tea is an exercise of her ability to make a cup of tea, an agent's raising her arm is an exercise of her ability to raise her arm. On this view, each ability has one action type as its exercise: in exercising an ability my action must be of the type that this ability is for. When I am simply raising my arm, I am exercising my ability to raise my arm—an ability for self-movement. When I am raising my hat, I am exercising a different ability.

On this single-track picture of abilities, the exercise of an ability to move my arm must be an action of moving my arm—since each ability has one corresponding action type as its exercise. So, to say that all human agents' transactions are exercises of bodily abilities is to say that all human agents' transactions are actions of moving the body. This appears to contradict the Materialist picture of transactions.

Adopting a multi-track picture of abilities, on which a single ability may have many action types as its exercise, changes this. It allows us to see how a single ability to move my arm can have two different action types as its exercise. On this view of abilities, it makes perfect sense to say that an agent's transacting with another object is an exercise of one or more of the agent's abilities to move her body. It can be the case that the ability I exercise when I raise my arm is the same ability that I exercise when I raise my hat. There is no separate ability required in order to raise my hat in addition to my ability to raise my arm. So, with this picture of abilities we have the beginnings of a way to understand the embodied nature of human agency without stepping on the toes of Ford's claims about transaction. The abilities exercised in other-movement are the same as those exercised in self-movement.

Of course, even without the mutual manifestation view of abilities, Ford need not deny that when I raise my hat I am also raising my arm—that when I take part in a transaction I am also engaging in self-movement. But on the pre-theoretical view of abilities he would have to accept that I am simultaneously exercising two abilities: the ability to move my arm and the ability to move my hat. He must accept that there is a new ability on the scene when I am transacting in addition to the ability being exercised when I am simply moving my body, even though both are exercised simultaneously in transaction. Although there is nothing contradictory in this picture, it 'separates out' the self-movement and the other-movement in a way that

echoes the splitting up of transaction that Ford is attempting to avoid. Further, it shows how self-movement occurs alongside transaction but offers less of an explanation of why the former is required for the latter. Seeing both actions as the exercise of the same ability ensures we are not drawn into separating out agents' transactions into two parts, and shows the importance of our status as embodied beings in our interactions with other objects.

3.3 Matching exercise types to abilities: a problem

We have seen how, on a multi-track picture of abilities, it is possible for my taking part in a transaction to be an exercise of an ability to move my body. But at this point we appear to come up against a serious objection to the idea that transactions are exercises of bodily abilities.

To introduce the problem, let us take someone's picking a lock as an example of a transaction. On the view I am proposing, the human agent's transacting with the lock is an exercise of the ability to move her hands. The issue with this is that many people have the ability to move their hands but cannot pick locks. It is possible that Jeremy and Janet both possess the ability to move their hands, but Janet can pick a lock whilst Jeremy cannot. It sounds wrong, then, to claim that picking a lock is an exercise of the ability to move one's hands. If it were, both Jeremy and Janet should be able to pick locks, because they both possess the ability to move their hands. It seems that in order to account for this difference between Jeremy and Janet, we should appeal to a further ability that Jeremy lacks: the ability to pick locks.

What we need at this point is a principle by which to decide which actions are exercises of which abilities. Adopting the right principle will show us why picking a lock's being an exercise of an ability to move one's hands is compatible with both Jeremy and Janet possessing the ability to move their hands whilst one of them cannot pick a lock. Such a principle will also guide us in deciding which transactions are the exercises of which bodily abilities: it will show us why picking a lock is an exercise of the ability to move one's hand rather than, say, the ability to move one's feet.

In the discussion of multi-track dispositions in Chapter 3 Section 3.4 we looked at principles that could be used in deciding which dispositions have which manifestations. We looked at Neil Williams' idea that when powers are 'clustered' they are identical, but did not settle on any one principle.[30] Here I will introduce a principle proposed by Don Locke, which can be applied both to abilities and dispositions.[31] Locke's approach to individuating abilities gives us the tools needed to show why the apparent problem above should not concern us—but some aspects of his principle will need modifying if they are to prove useful as a guide to which actions are exercises of which abilities.

Locke principle is as follows:

> *If those features from an agent's constitution and background which bring it about that he standardly succeeds in doing x are also sufficient to bring it about that he standardly succeeds in doing y, then we are dealing here not with two separate abilities, but with the one ability, to do x and do y.*[32]

The idea behind this principle is that an agent's possessing an ability to do x is a matter of the agent having characteristics that allow her to reliably do x. Laura's running up a hill is dependent on various of her characteristics—such as her limbs and muscles being in working order and strong enough to propel her up a gradient. It is in virtue of her having these characteristics that she has the ability to run up a hill, and it is this ability she exercises when she runs up a hill. When another action—such as her running down a hill—seems to be dependent on the same characteristics, we should not then propose that Laura's running up a hill and running down a hill are exercises of two separate abilities. She is making use of the same characteristics in both cases, and so we should see her as exercising the same ability in both cases.

Locke's principle may seem to run into a problem, in that it individuates abilities in a more fine-grained way than he intends. The characteristics necessary for running up a hill and running down a hill must be different, since some agents can run downhill but not uphill—perhaps their fitness level is such that they cannot manage to run on any upward gradient but can manage a downward gradient with little trouble. So,

30 (Williams 2011)

31 (Locke 1973)

32 Ibid. p. 187. Quoted by Will Small (2016).

given that these two actions require different characteristics, Locke's principle gives the result that running uphill and running downhill are exercises of separate abilities. It seems that the same issue can arise for any two actions: we can always imagine a case in which one agent can do both but another agent can only do one of the two. Swimming 500 metres and swimming 510 metres; lifting a teacup and lifting a mug; cutting a ribbon and cutting string: in each of these cases it is likely that some agents can do one but not the other—so it appears we must say that each pair of actions involves two separate abilities being exercised.

The problem, however, arises from a misreading of Locke's principle. Seeing how Locke avoids this problem will show us how his approach is helpful to us in overcoming the challenge to the idea that transactions are exercises of bodily abilities.[33] Locke is not aiming to provide an answer to the question of whether the ability run up a hill is identical to the ability to run down a hill, but rather trying to answer the question of whether a particular agent's abilities to run up a hill or run down a hill are different or are in fact identical. The characteristics that ensure that I have the ability to cycle 10.1km are also what ensure that I am able to cycle 10km—so Locke would say that my cycling 10km and cycling 10.1km are both exercises of my more general ability to cycle. This is true even though someone else might have the ability to cycle 10km but lack the ability to cycle 10.1km. All this shows is that my ability to cycle is different to theirs, and I can exercise my ability in ways that they cannot.

When we see abilities as specific to individuals rather than specified generally it seems less strange to think that human agents' transactions with objects are exercises of their abilities to move their bodies. The possible objection was that agents often possess the same bodily abilities but can do different things—Jeremy and Janet both possess the ability to move their hands but only Janet can pick a lock. On the view given by Locke's principle, we can see why this line of thought is wrong—and we can thereby see why Janet's picking a lock can in fact be an exercise of her ability to move her hands. Jeremy's ability to move his hands has a different nature from Janet's ability to move her hands. Janet has learned to exercise her ability to move her hands in a way that Jeremy cannot—she has learned to pick locks. The fact that Janet's picking a lock is an exercise of her ability to move her hands does not mean

33 Locke's method of avoiding the problem is discussed by Will Small (2016).

that anyone with the ability to move their hands must also have the ability to pick locks.

Taking on board Locke's insight that we should think of abilities as specific to individuals has helped us avoid this objection to the view that transactions are exercises of bodily abilities. Now, we can look at how Locke's principle guides us in deciding which actions are exercises of which abilities.

3.4 Which actions are exercises of which abilities?

Locke argues that in order to decide whether an agent's doing x and doing y are exercises of the same ability, we should ask which features bring it about that an agent standardly succeeds in doing x and which bring it about that an agent standardly succeeds in doing y, and then look at whether these features are the same. Given that an agent's bodily characteristics are crucial in allowing her interaction with the world, it seems that Locke's principle converges on bodily agency. The characteristics that bring it about that I succeed in tying a shoelace, opening a bag of crisps or clicking my fingers are characteristics such as dexterity with my hands. Locke's principle, then, guides us toward the idea that these actions are all exercises of the same ability: the ability to move my hands.

There are problems with Locke's principle and the way it is applied here that suggest it is in need of modification, however. It is not clear exactly what we are supposed to be looking for when we ask which 'features from an agent's constitution and background' bring it about that the agent succeeds in doing x. An agent's bodily characteristics are evidently necessary for the agent's success in acting, but it might be claimed that the agent's knowledge, and character traits such as perseverance, or mental capacities such as rationality, are also necessary. Such a wide range of features are necessary for agents' doing the things they do that Locke's principle does not provide us with a useful guide in deciding whether the features in any two cases are the same.

Even when we focus on an agent's bodily characteristics, there is a wide range to choose from. I mentioned dexterity, but we could also bring in features such as the

length of the agent's arms, the strength of certain muscles and so on. Locke's principle leaves us with too many features to look at, meaning that agreement about which actions are the exercises of which abilities will be difficult to come by if we follow his approach.

Instead, we need a principle that taps into the way we perceive and intuitively understand the way agents interact with others. We should ask, not which features of an agent bring it about that the agent succeeds in doing what she does, but *how the agent is doing what she is doing*. If how the agent is doing x is the same as how the agent is doing y, then the agent is exercising the same ability in both cases.

We first need to see how the question 'how is the agent doing x?' should be understood. The question is not intended to fix on an agent's practical reasoning or the teleology of an action, but is meant to elicit a more detailed description of the process of transaction that is taking place between the agent and the object she is acting on. For example, the answer to 'How is Jimmy picking up the cup?' might be 'He is picking it up with his hand', 'He is picking it up with his elbows', or 'He is picking it up with his teeth'. Answering the question gives us more information about the transaction than simply asking what the agent did—which could be answered with 'He picked up the cup'—and highlights the fact that the same action type may involve very different abilities on different occasions.

A similar idea is mentioned by Will Small.[34] In discussing the best way to convey the thought that human agents' embodied nature is central to our agency, he suggests that instead of saying 'She raised the glass' it would be better to use phrases such as 'She right-arm-raised the glass' or 'She left-arm-raised the glass.' This conveys the nature of the process that is taking place better than simply saying that the agent raised the glass. It also allows us to be clearer about which of the agent's abilities she is exercising in doing what she does.

This view gives the result that the agent is exercising the same ability when she lifts a cup with her hand and when she lifts a flowerpot with her hand—but is exercising an altogether different ability when she lifts a cup with her mouth. By asking not just

34 (Small [forthcoming]) p. 20

what the agent is doing but how she is doing it, we can better understand the powers that she is exercising when she takes part in transactions.

One of the advantages of appealing to how an agent is doing x, rather than asking which features and characteristics are responsible for her successfully doing x, is that we can specify more clearly what kind of answers we seek. We have seen that there are a wide range of features and characteristics that we could use to answer Locke's question, making agreement on whether two actions are exercises of the same ability unlikely. The question of how an agent is doing what she is doing has a prescribed sense, and can be answered simply through our perceiving agents' actions and giving a description of the process that is taking place.

This is not to say that the description required to answer the 'how' question is a natural one. As noted by Anscombe, when we see people acting we naturally perceive what they are doing intentionally:

> *I am sitting in a chair writing, and anyone grown to the age of reason in the same world would know this as soon as he saw me, and in general it would be his first account of what I was doing; if this were something he arrived at with difficulty, and what he knew straight off were precisely how I was affecting the acoustic properties of the room (to me a very recondite piece of information), then communication between us would be rather severely impaired.*[35]

The aim of the 'how' question, however, is to ask for additional information about the transaction taking place that is not contained in our descriptions of action types. It is not information that we would normally look for, which is why the question is needed, but it is information that we naturally pick up on and have no trouble giving—unlike information about how an agent is affecting the acoustic properties of a room.

The principle I have proposed is not immune to being interpreted in different ways, however, and does rely on our having an intuitive understanding of the appropriate level of specificity with which to describe how an agent is doing what she is doing. Rather than saying that she raised the cup with her hand we might say that she raised it by pinching three fingers and a thumb, or that she picked it up by using her whole body to balance herself and ensure she did not fall over. If we do not answer the

35 (Anscombe 1957) p. 8

question to the right level of specificity, the principle is in danger of becoming useless. This is not, however, a problem that arises from the idea that we should ask how an agent does something—rather, it is a problem common to many attempts to give principles of individuation. It does not, then, seem like a problem with which we should be overly concerned.

We can demonstrate that a similar problem arises in both of the principles we have looked at so far. Recall Neil Williams' principle, discussed in Chapter 3 Section 3.3: 'for event types φ and ϕ, where φ-ing ≠ ϕ-ing, the power to ϕ and the power to φ are clustered just in case the class of objects with the power to ϕ is identical with the class of objects with the power to φ.'[36] Williams argues that when powers are 'clustered', we have a reason to think that they are in fact identical. Williams' principle falls down if our application of it becomes too fine-grained or specific.

If we have a sufficiently fine-grained view of powers, then it seems that the class of objects possessing the power to ϕ will very rarely be identical to the class of objects with a power to φ. We would have to say that an elastic band's power to stretch to 15cm is different from the elastic band's power to stretch 15.3cm, because there are some elastic bands whose stretching limit is 15cm and so lack the power to stretch 15.3cm.

The same occurs with Locke's principle. If we are detailed enough in our listing of what is required for an agent's doing something, we would end up with the conclusion that every action is made possible by different characteristics from every other action. Perhaps Laura's running 5km is made possible by some mental fortitude or tolerance to pain that she need not to draw on when she runs 4km. Perhaps Ben's tying his shoelace and tying a bow in a ribbon require knowledge of the process for tying different kinds of knots. We can always find some difference in what is required. The results of Locke's principle depend on how meticulous or fine-grained we are in our descriptions, and on how detailed our accounting of the characteristics required for that action is.

These are not reasons to discard these principles, but merely reasons to see that such principles should be taken as guides rather than hard and fast rules. Similarly, the

[36] (Williams 2011) p. 582

principle I have suggested cannot be of use if applied in a vacuum, but relies on a common intuitive understanding of the right level of specificity, informed by our normal perceptions of what it is that people are doing. Only then can it be a useful guide to how we should decide which abilities an agent is exercising.

The final point to note about this view on which transactions are exercises of bodily abilities is that it need not diminish the importance of human agents' mental abilities. Most of our actions involve not only bodily powers but also require mental powers: my picking the lock requires that I know something about the nature of locks and their workings, and my walking to the shop requires that I know where the shop is. The aim of this view is to show in what sense our embodied nature is central to our agency: we exercise our bodily abilities in getting things done. But mental capacities play a crucial role in allowing us to exercise our bodily abilities in different ways. Learning about locks allows me to exercise my ability to move my hands in a new way, in picking the lock.

3.5 Self-movement and transaction

Ford distinguishes between self-movement and other-movement but remains silent on the role self-movement plays in human agents' transactions. It is very unlikely that Ford would want to deny the fact that self-movement plays a part in everything we do. After all, we are embodied beings and cannot bypass our embodied nature even in cases where we directly act on other objects. It is more likely that Ford's silence on this issue stems from a wariness of being led back to a picture on which transactions are split up into two parts. But through developing a multi-track view of abilities and a principle for their individuation, we have shown in what sense self-movement plays a role in all of our actions without resorting to splitting up transactions into self-movements and their further effects.

A great many human actions, including transactions, are exercises of bodily abilities. The key to this initially implausible-sounding claim not standing in contradiction to Ford's view that agents 'do what happens' is the fact that the exercises of bodily abilities need not be actions of bodily movement. An agent's exercise of an ability to move her hand need not be an action of moving her hand—it may be an action of

raising her hat, flipping a coin, pressing a switch and much else. We should decide which actions are exercises of which bodily abilities by paying attention to *how* an agent is doing what she is doing, through giving a description of the physical process the agent is engaging in that goes beyond the information given in our descriptions of action types. This tells us which of the agent's bodily powers she is exercising in a transaction.

In the course of this discussion we have come to further conclusions about the characteristics of abilities. We have seen that abilities of the same type may have different exercises when possessed by different agents. It has become clear that we cannot specify how the ability to move one's hand may be exercised—only how the Janet's ability to move her hand may be exercised or how Jeremy's ability to move his hand may be exercised. Agents may gain characteristics and knowledge that allows them to exercise their abilities in different ways.

Conclusion

We should accept that human agents take part in transactions. But we must be careful that in accepting this we do not forget the role of self-movement in all of our actions, and thereby lose the sense that human are embodied agents. What Ford fails to do, we have seen, is accommodate our notion that a bodily ability is in play in all of our actions.

The account of abilities I have proposed taps into the fact that we perceive and experience the bodily processes that constitute our transactions at a level of detail not specified by action type descriptions. I have proposed that we use this understanding to individuate abilities: if the answers 'how is Jenny doing X?' and 'how is Jenny doing Y?' converge on the same type of bodily process, Jenny's doing X and Jenny's doing Y are exercises of the same bodily ability. With this view it becomes clear that the abilities exercised in self-movement are the same as those exercised in other-movement.

Chapter 6

Abilities and Non-human Animals

The previous chapter developed a picture of the bodily abilities exercised in human agency. The current chapter is concerned with the question of whether non-human animals, like human agents, possess abilities to move their bodies.

It is clear that at least some non-human animals possess abilities, but the more difficult aspect of this question is the issue of where the line should be drawn between those animals that possess abilities and those that only possess dispositions. It seems that there must be some cut-off point for the possession of abilities: whilst it is clear that, say, a chimpanzee possesses abilities, it is less clear whether a slug or a flea possess abilities or only possesses dispositions.

Answering this question will present the opportunity to flesh out the picture of bodily abilities developed in Chapter 5. To determine where the lines should be drawn, we must ask which characteristics are prerequisites of ability and look for evidence of these characteristics in non-human animals. I will argue that possessing abilities requires only the possession of a conscious perspective, and does not require the possession of higher-level mental capacities distinctive of human agency.

1.

Do Non-human Animals Possess Abilities?

1.1 Some non-human animals possess abilities

It will appear obvious to most people (including philosophers and scientists researching animal behaviour[1]) that at least some non-human animals possess abilities to move their bodies. The view that our everyday conception of animals as

1 Cf. (Allen & Beckoff 1999) pp. 1-21

agents is 'anthropomorphic' and that animals are merely mechanistic systems is not mainstream.

This is to some extent due to the widespread acceptance that, in the case of most species, mechanistic explanations of animal behaviour in terms of stimulus-response systems, motor patterns and so on would have to be incredibly—often impossibly—complex in order to account for the flexibility and adaptiveness of (at least some) animal behaviour. This flexibility and adaptiveness was previously ignored in behaviourist and mechanistic accounts of animal behaviour—but advances in techniques for observing and experimenting on animals have drawn these features into sharper focus.[2]

Nevertheless, it is still worth illustrating the mechanistic view—from which arises the idea of animals lacking abilities—and comparing it to our pre-theoretical picture of agency. Looking at an example of a view on which animals only possess dispositions and lack abilities will show us what we are claiming of the animals we decide do not have the ability to move their bodies.

Comparing two descriptions of the same range of animal behaviour from the scientific literature will illustrate how different the view on which animals lack abilities is from our pre-theoretical view of animal agency. The example I will focus on is the mating behaviour of sticklebacks.

Part of the mating ritual of sticklebacks involves the male putting on a display for the female. Here is a description from the 19th century scientist Robert Warington:

> *He darts around her in every direction, then to his accumulated materials for the nest, then back again in an instant; and as she does not advance, he endeavors to push her with his snout, and then tries to pull her by the tail and side-spine to the nest.*[3]

Compare this description, which chimes with the way one might talk about this behaviour on observing it in the course of everyday life, with a description by ethologist Nikolaas Tinbergen of the same stickleback behaviour:

---

2 (Allen & Beckoff 1999) pp. 1-21. Allen and Bekoff note that unexpected behaviours by animal subjects in experiments were often dismissed in recorded results as statistical outliers.

3 (Warington 1852) p. 277

> *Each reaction of either male or female is released by the preceding reaction of the partner […] The male's first reaction, the zigzag dance, is dependent on a visual stimulus from the female. The female reacts to the red color of the male and to his zigzag dance by swimming right towards him. This movement induces the male to turn round and to swim rapidly to the nest. This in turn induces the female to follow him, thereby stimulating the male to point his head into the entrance. His behaviour now releases the female's reaction; she enters the nest […]This again releases the quivering reaction in the male that induces spawning. The presence of fresh eggs in the nest makes the male fertilise them.*[4]

In the technical vocabulary it uses, this second description is instantly distant from our everyday thought about animal agency: Tinbergen uses terms such as 'visual stimulus', 'stimulate', 'release' and 'induce'. However, other language is much the same as a 'layman's' description of animal agency. Tinbergen uses verbs that attribute agency to the fish—they 'dance', 'swim', 'follow', 'turn round'. Although looking at the vocabulary can help us start to see the difference between the two views of the sticklebacks' agency, we must look deeper to bring out the real differences between the two pictures.

For Tinbergen and other ethologists who share his approach to understanding animal agency, stimulus-response mechanisms are central to animal behaviour. Hence Tinbergen's description of the stickleback being 'induced' or 'stimulated' to move in different ways by particular stimuli. Tinbergen developed a theory of animal behaviour on which motivational impulses build up until they are released by stimuli in the environment, prompting a fixed pattern of behaviour.[5] This is why he sees each stickleback's behaviour being 'released' by movements of the other.

Those who hold this view do not explicitly mention dispositions—but the stimulus-response mechanisms they hold to be the drivers of animal agency are clearly recognisable as grounding dispositions of the kind discussed in earlier chapters. On Tinbergen's picture, animals' movements are the result of dispositions manifesting in response to stimulus conditions—these animals are not in charge of their bodies in the way that comes along with exercising abilities. On this description, the stickleback's behaviour becomes rigid, and what seemed like a flexible interaction

4 (Tinbergen 1951) p. 47

5 (Tinbergen 1951)

between two fish becomes a chain of reactions, each manifestation of one fish's disposition triggering the manifestation of the other fish's disposition.

Where this stimulus-response view of animal agency was once more common, it has been replaced by a view of a continuum between those animals whose behaviour is rooted in instinct and stimulus-response mechanisms (i.e. their behaviour is the manifestation of a disposition) and animals whose behaviour is 'stimulus-free' (i.e. their behaviour is the exercise of an ability).[6] A version of this continuum may exist in the range of behaviours of a single animal. For example, a dog may exhibit some behaviours that are based in instinct and are manifestations of dispositions, such as licking its lips when it sees food, and other behaviours that are not tied to particular stimuli and are exercises of abilities, such as searching around a whole house until it finds its favourite toy.

It is clear, then, that at least some non-human animals do possess abilities. The more interesting question to ask is *which* animals possess abilities—where exactly we should draw the line between those animals that possess abilities and those that do not. It will take the remainder of the chapter to answer this more difficult question.

1.2 Does possessing ability require a capacity for intentional action?

The question of which animals possess the ability to move their bodies and which animals do not cannot be adequately answered by consulting our pre-reflective intuitions. I have no strong pre-reflective intuitions about whether or not an ant has the ability to move its body, for example. The way to arrive at an answer will be to look more closely at our concept of ability and the conditions under which we apply it.

The way forward is to decide which characteristics an animal must possess in order for the attribution of an ability to seem plausible. In cases where it is unclear whether it is correct to attribute an ability, we can ask whether the animal possesses these characteristics.

6 A range of such views is discussed in (Allen & Beckoff 1999).

When we think about what is required for the exercise of an ability to move one's body, one thing tends to stand out. It is a mainstream view among philosophers of action that all actions must be intentional, and when we think about human agency it is difficult to come up with a case where nothing is done intentionally by the agent yet we would still want to say that the agent exercised an ability.

I may unintentionally knock a glass from a table when I get up from my seat—here we still want to say that I am exercising an ability when I am knocking over the glass, but that is only because I am doing something intentionally: getting up from my seat. The ability I am exercising is the ability to get up from a seat, not the ability to knock over a glass. Where an agent does something but nothing is done intentionally—flinches, falls over or sneezes—this appears to be a manifestation of a disposition. It is a movement in response to some stimulus, rather than an ability's being exercised by the agent.

This sets a bar for the possession of abilities—but it sets a very high bar. Many believe that doing something intentionally involves doing that thing for a reason.[7] Acting intentionally is taken to involve capacities for rationality, practical reasoning and reflective consideration of one's plans for action that are beyond many, if not all, animals. If, as appears to be the case, exercising an ability must be intentional, this would exclude these animals from possessing abilities.

Given that in the previous section we saw how implausible it would be to deny that many non-human animals possess abilities, it seems the most reasonable thing to conclude in response to learning that animals cannot possess intentions is that intentions are not required for ability. If many animals lack the capacity for intention, but it would be implausible to deny that these animals lack abilities, we must conclude that intentions are not required for ability.

As straightforward as this argument seems, it will not do to leave our discussion of intentions there. If we want to conclude, based on our assumption that many animals possess abilities, that intention is not required for the possession of ability, we need to do more to show *how* ability comes apart from intention. We have seen how in nearly all cases of human action, abilities must be exercised intentionally. To

7 Cf. (Mele & Moser 1994); (Bratman 1987)

satisfactorily conclude that animals' exercises of ability need not be intentional, we must say how and why the animal case differs so much from the human case.

In Chapter 4 Section 3.4 we looked at sub-intentional actions. Sub-intentional actions like scratching my head while thinking or absent-mindedly tapping on the desk are things I am often not aware I am doing: I may only realise I am tapping on the desk when the person sitting next to me asks me to be quiet. Because of this lack of awareness, it is difficult to say that these actions are goal-directed or done for a reason—and they certainly do not seem to be intentional. However, we readily ascribe these behaviours to ourselves as things that we do, rather than putting them into the category of automatic movements like breathing or digesting. We take it that we are exercising our abilities when we do these things—these movements are not just the manifestations of dispositions as blinking or sneezing are.

We might be tempted to try to use sub-intentional actions to understand how animals' abilities need not be exercised intentionally. If animals' actions are like sub-intentional actions, this would explain why animals can exercise abilities without having the capacity to do anything intentionally. Unfortunately, this line of thinking will not work. Examples of sub-intentional action may show that not all exercises of an ability must be intentional. However, it is a further step to say that an animal that lacks the ability to do things intentionally could still exercise an ability to move its body.

Sub-intentional actions make up a proportion of the things that human agents do, but many of the things that human agents do are intentional. It looks as if in order for it to be the case that an animal could possess an ability to move its body without possessing the capacity to exercise this ability intentionally, *all* of that animal's actions would need to be sub-intentional. This is very different from the human case. So, further consideration is required to decide whether this is possible. There are two reasons why it is implausible to claim that animal agents only act sub-intentionally.

The first reason is that the idea that animal agents only act sub-intentionally does not fit in with our everyday thought about animal agency. A cat carefully stalking a bird, waiting for the right moment to strike, seems to us to be aware of what it is doing, and acting purposively. To claim that all of this animal's actions are sub-intentional,

one would have to claim that the cat somehow has no awareness whatsoever of what it is doing as it follows the bird, and that its movements just spring from some impulse in the cat with little direction or purpose—just as my absent-mindedly tapping my fingers on the desk. This is to radically rethink the behaviour of an animal that seems to us to be carefully focusing on a task.

The second reason is that it appears impossible for an agent to have the capacity to exercise an ability sub-intentionally unless it also has the capacity to intentionally stop or continue doing what it is doing. As noted in Chapter 4's discussion, key to the thought that sub-intentional actions are *actions* rather than automatic movements is the fact that they are behaviours we can become aware of and stop or continue intentionally. When I am tapping my finger on the desk and my neighbour points it out to me, I become aware of what I am doing and can continue doing it or decide to stop.[8]

An animal that lacks the capacity to act intentionally cannot intentionally continue or stop doing what it was doing sub-intentionally. This explains to some extent why the above description of the cat stalking the bird appears so strange. If what the cat is doing is not intentional and does not feature in its awareness, and in addition there is no way for the cat to become aware of these movements and continue making them intentionally nor stop making them, then it does not seem that the cat is acting at all. The cat's stalking its prey has the same status as the cat's sneezing or digesting food.[9]

This shows that sub-intentional actions cannot be invoked to explain why animals need not exercise abilities intentionally while humans must. A better explanation is required if we are to conclude that intention is not necessary for the exercise of ability.

---

8 There are arguably some abilities that cannot be exercised intentionally, such as the ability to doodle or the ability to daydream. These are actions that are inherently sub-intentional. But, whilst one cannot intentionally continue doodling or daydreaming once one has realised what one is doing, one *can* become aware of and intentionally stop what one is doing.

9 Perhaps the suggestion that all animal actions are sub-intentional could be made more workable by adopting Helen Steward's suggestion, discussed in Chapter 4 Section 3.4, that there is some 'lower' form of awareness still present in the case of sub-intentional action (Steward 2013) pp. 698-699. Whether this would avoid the implausible consequence that the cat is unaware of its stalking its prey, will depend on how the idea of a lesser degree of awareness is fleshed out.

Fortunately, a better explanation is available. It is because of human agents' capacities for rational thought that there is a link between ability and intention in human agency. The existence of this link in the case of human agency is compatible with the lack of such a link in the case of non-human animals.

Human capacities for rational thought make intentional action possible, but rational thought is also in some senses 'inescapable' for human agents: it colours everything we do. It means that in nearly all cases where we exercise ability we do so intentionally, despite the fact that there is nothing in the nature of ability that requires that these exercises are intentional.

Our rational capacities allow us to make plans for the future, reason about our course of action and say what our reasons for doing things are. But the fact that we cannot 'switch off' these capacities means that very rarely can we act to achieve something without doing so intentionally.[10] This stands in contrast to the actions of animals that lack rational capacities—these animals can exercise their abilities without doing so intentionally, because they do not possess the rational capacities that infuse all human exercises of ability and make them intentional. An animal may do things that are goal-directed—run across a field in pursuit of its prey—without doing so intentionally.

This point can be illustrated by looking at a distinction drawn by Anthony Kenny between human and animal desires.[11] Kenny argues that human and animal desires are two separate species of the same genus of 'wants'. Animal desires—for example, the desire for food or the desire to hide from danger—seek immediate satisfaction, and are felt more or less continuously until they are satisfied or otherwise disappear. Human desires may be for something in the future, and need not be consciously felt at all times. The reason for this difference, Kenny argues, is that all human desires are 'coloured and modified by our possession of language'.[12] It is humans' conceptual abilities that give our desires such a different character to those of animals.

Crucially, Kenny claims that because of our conceptual abilities, human agents are *unable* to have the kind of 'pure' felt desire that animals possess. Our conceptual

10 The exception to this is the case of sub-intentional action.

11 (Kenny 1989) Ch.3

12 Ibid. p. 36

abilities mean that our desires will necessarily have a different character to animals'. Even the human desire for food is 'coloured' by our abilities to think about the future: 'as my hunger becomes progressively more uncomfortable I keep thinking "still an hour and twenty minutes to go before dinner", using concepts that are well beyond the competence of a hungry dog.'[13]

I do not believe that there are two separate species of ability, one possessed by animals and the other possessed by humans. But Kenny's point that humans' conceptual abilities 'colour' their desires lends support to my claim that humans' rational capacities affect the nature of their exercising their abilities. This is why, when humans exercise abilities, they must do so intentionally, whilst animals who lack the rational capacities of humans exercise abilities without doing so intentionally. It is not a feature of ability itself that it must always be exercised intentionally—it is just that, in the human case, it is always exercised intentionally because it is always exercised by beings with certain rational capacities that affect the nature of their exercising of their abilities.

This provides an explanation of why intention is not necessary for ability, despite its being the case that human agents must nearly always exercise abilities intentionally. Unfortunately, this also leads us back where to we started: we have yet to find the characteristic that *is* a prerequisite for the possession of abilities. So we have not yet made much progress towards finding a way to decide which animals possess abilities and which do not.

13 (Kenny 1989) p. 36

2.

Ability and Conscious Perspective

2.1 Possessing abilities requires a conscious perspective

I will suggest that in order to possess and exercise abilities, an agent must possess phenomenal consciousness. Phenomenal consciousness imparts the agent with a single animal-level perspective on the world.[14] Such a perspective unifies the work of the various bodily mechanisms involved in action, and means that it makes sense to see certain movements as coming from the agent as a whole rather than simply from combinations of sub-personal processes. I will first say more about what I mean by phenomenal consciousness and the perspective it involves, before showing how possessing such a perspective is an indicator of the possession of abilities.

Simply put, possessing phenomenal consciousness is a matter of having subjective experiences that are like something to one: seeing the colour green, feeling the cold, hearing birdsong. These subjective experiences by their nature can only be had from a point of view. It is in this way that phenomenal consciousness involves having a perspective—a continuous, singular viewpoint from which experiences are had. This goes for all animals, even if the nature of the things they can experience is wildly different and unimaginable from the human perspective, as in Nagel's famous example of the bat's experience of the world through sonar.[15] These experiences must still be had from a single perspective.

There are many different theories of phenomenal consciousness, but there is not space here to enter this debate. I will not be using the understanding of phenomenal consciousness that is developed on higher-order theories of consciousness. On these theories, what makes a state conscious is its being available to higher-order thought. For example, Peter Carruthers argues that in order for a perceptual state to be

14 My use of the term 'animal-level perspective' is intended to apply to the perspectives of human and non-human animals.

15 (Nagel 1974). There is a difference between my view and Nagel's: Nagel's view appears to be that perspectives are specific to kinds of animal, whilst I am using 'perspective' to refer to something specific to individuals.

phenomenally conscious the agent must hold higher-order thoughts about that state.[16] Since many non-human animals are incapable of higher-order thought, this view from the beginning rules out that many animals possess consciousness or abilities.

I believe that an opposing approach to phenomenal consciousness, a 'first-order' theory, provides a better understanding of phenomenal consciousness and the importance of a conscious perspective for possessing and exercising abilities.[17] On this theory, what makes a state conscious is its being available to influence an agent's first-order thought—the agent need not have thoughts about its conscious experiences. An example of an experience affecting first order thought is the experience of seeing an apple on the table leading me to believe there is an apple on the table. My thought is not about my experience, but is about the apple itself.

On this view, the prerequisites for possession of phenomenal consciousness are not too demanding—no higher-order thought is needed—and at the same time phenomenal consciousness is by no means epiphenomenal.

Having conscious experiences is a matter of taking in information about one's environment that makes a difference to one on the personal, or animal, level: an experience is something had by the *agent*, not by some part of the agent or some sub-personal mechanism within the agent. If the agent has a perspective, there is at least some information that is not gained through sub-personal processes, but rather is gained by the agent as a whole. It is this unifying role of the perspective that is important for ability.

In Chapter 4 we looked at Steward's attempts to say how it is that the agent, rather than sub-personal mechanisms within the agent, is in control of the body during an action. We saw that instead of trying to show that the agent has a role in overseeing the workings of sub-personal mechanisms that are separate to the agent, we should simply see the work of these mechanisms as part of the agent's exercising her abilities. The idea that a conscious perspective is linked to the possession of abilities reinforces this idea. Although sub-personal mechanisms are at work to some extent in all of an agent's actions, the agent's possessing a perspective entails that there is

16 (Carruthers 2000)

17 Cf. (Tye 1995)

some level on which the *agent*, rather than a collection of parts of the agent, is interacting with its environment and exercising its abilities.

The importance of a conscious perspective for the possession of abilities is cemented when we see that it is also necessary for the agent's possessing goals, which will be discussed in the next section. First, it is important at this point to make a clarification about the notion of a conscious perspective.

Having a conscious *experience*, such as the experience of pain, is different from possessing a conscious *perspective*. An experience of pain accompanies my pulling my hand away after accidentally touching a hot pan—this sensation accompanies a reflex action. I am using 'conscious perspective' not to refer to the experience of a sensation like pain in isolation—rather, I am referring to a richer subjectivity that incorporates many different experiences. A conscious perspective involves what is sometimes characterised as 'the unity of consciousness', and in a helpful discussion by Gerald Edelman is described as what allows an animal to 'integrate perceptual and motor events together with memory to construct a multimodal scene in the present.'[18] A conscious perspective is a perspective on the world, more than an isolated experience such as the sensation of pain.

Does this mean that an animal could have 'mere' conscious experience without having a conscious perspective? At first look this seems conceivable: it seems we can imagine a being that, say, only has the capacity to feel pain and lacks any other experience. But on closer inspection it becomes clear that on our understanding of consciousness, experiences are never had in isolation and always form part of a conscious perspective.

Let us visit again the case of pain when burning my hand on the pan. I touch the pan, then I reflexively pull back—a movement that is the manifestation of a disposition—and this movement is accompanied by pain. The pain does not play any role in my automatic reaction of moving my hand away from the pan. However, as an on-going sensation felt in my finger my pain *does* play a role in affecting my other behaviour: I will refrain from touching the pan again, will be more likely to run my hand under cold water, and so on. The pain plays these different roles because it

18 (Edelman 2003) p. 5521

is part of a conscious perspective that involves other experiences—a perspective from which I have a picture of what has happened. My pain is not an isolated experience but is had as part of a conscious perspective.

The same seems to be true in the case of other animals—evidenced by the fact that what researchers look for as indicators of pain is a variety of different behaviours rather than a reflex movement away from a harmful stimulus. Michael Tye notes that we take an experience of pain by its very nature to play a role in various behaviours:

> *One of the things that is striking about the experience of pain is the effect it has above and beyond simple withdrawal from the noxious stimulus. For example, mammals feeling pain may stop moving, limp, protect the damaged part, show irritability, become aggressive if approached, feed less, and decrease their sexual activity.*[19]

Again, we look for this behavioural evidence because we take it that experiences like the sensation of pain influence behaviour as part of a richer conscious perspective. They do not simply accompany reflex movements. This is why reflex movements such as drawing away from a harmful stimulus are not taken as indicators that an animal is undergoing an experience unless other behaviours such as those above are also present.

So, although possessing a conscious perspective and having a conscious experience are different, it seems that having a conscious experience is in most cases sufficient for having a conscious perspective. It may seem possible to imagine a being that is only capable of experiencing an isolated sensation such as pain. But in this case, the experience would play no role in the being's behaviour—based on what we know about conscious experience, it seems that this isolated experience could only accompany reflex action. It is therefore difficult to explain why such a being would have evolved with these traits, and difficult to ever gain evidence for the existence of such experience. Given what we know about conscious experiences, we should take it that they are not isolated but are always had as part of a richer conscious perspective.

19 (Tye 2016) p. 137

2.2 Conscious perspective and goal-directed behaviour

In this section I will show that a conscious perspective is essential for an agent's possession of goals. Without this single, unified, perspective, the only purpose we can attribute to an agent's behaviour we must derive from our understanding of natural selection and the idea that all biological organisms' ultimate aim is to survive and reproduce. Goals assigned on the basis of our understanding of that being's evolutionary history are not goals of the agent, and are compatible with the agent's merely having dispositions rather than having abilities.

The distinction between two senses of 'goal-directed'—one referring to goals of the agent, one referring to goals derived from natural selection—is an important one. After clarifying this distinction, we will be able to see how the former sense of goal-directedness is central to the exercise of ability, and how it is linked with the agent's possessing a conscious perspective—further reinforcing my claim that the possession of a conscious perspective is what we should be looking for when deciding which animals possess abilities rather than only possessing dispositions.

The sense of 'goal-directed' that is not connected to the exercise of ability I will term 'derivative'. In most cases, attributions of 'derivative' goal-directedness are based on biological function, understood through natural selection. These attributions derive from our understanding of an agent's selectional history.

We can understand a biological entity's movements as purposive based on the traits for which that entity was naturally selected. For example, we may say that a human heart's pumping movements are goal-directed: the purpose of these movements is to keep blood flowing around the body. The human heart's movements are goal-directed in the sense that the heart has been naturally selected due to a tendency to bring about certain results, which these movements help to achieve.

A feature of these 'derivative' goals is that they are general: since we assign them based on naturally selected traits, any instance of a species or type of biological thing will have the same set of these 'derivative' goals. The movements of all human hearts have the same purpose, because all human hearts are things that have emerged from a process of natural selection that has favoured certain traits.

To see how this sense of goal-directedness applies to other cases, take another example: a sunflower twisting round to follow the sun's path through the sky. We might say the sunflower is following the sun with the goal of absorbing more sunlight—but if these movements are goal-directed, they are goal-directed derivatively. We understand these movements as goal-directed because they are movements which promote survival and are naturally selected for. Since its movements are derivatively goal-directed, the *sunflower* itself does not have a goal.

It is also the case that some human behaviours are goal-directed derivatively, rather than goal-directed by the agent. When I touch a hot pan and reflexively draw my hand back, there is a sense in which this movement is purposive—but only when understood in terms of natural selection. This reflex has survival value in that it makes sure my hand is immediately moved away from whatever is causing pain, and so receives less damage. The purpose of my movement, then, is to remove my hand from potentially being damaged. But this is not a goal that *I* have—from my perspective, my reflexive movement of my hand is something that just happens, and is not done in pursuit of a goal I have at all.

From looking at this last example we can see that 'derivative' goal-directedness of behaviour is compatible with the behaviour simply being the manifestation of a disposition. All that is required is that the movement has some purpose derived from natural selection—and in many cases the movements to which we can assign this kind of purpose are those triggered by certain stimuli. Tinbergen could, and likely did, accept that the male stickleback's mating behaviour has the goal of attracting a female—even though he saw it as simply a reflex reaction triggered by certain stimuli. If he saw the stickleback as having a goal, he saw it as having some 'derivative' goal—a purpose understood in terms of survival value and natural selection.

The second sense of goal-directedness and possessing a goal *does* imply that the behaviour is the exercise of an agent's ability. Unlike derivative goals, these are goals possessed by the agent.

In human behaviour we have many examples of goal-directedness in this second sense. When we see a person running to catch a bus, opening a gate or picking up something from the floor we see them as engaging in goal-directed behaviour, but

we do not see this behaviour through the lens of natural selection. We see humans as acting in order to achieve goals not because we take it that these behaviours will aid their survival, but because we take it that they are individuals in control of what they do, and are doing things in order to achieve their goals.

To attribute these goals, we do not need to look outside the agent. Contrast my picking a penny up from the floor with the reflex movement of drawing my hand away from a hot pan. We can see that both movements have a purpose, but only the former has a purpose that is *mine*. I am doing something in order to achieve my goal of picking up the penny. In the latter case, the movement does achieve a goal—but it is not a goal that I possess, and we must look to non-agential explanations in terms of natural selection to attribute such a goal to the movement.

What makes it possible for the agent to possess goals in the second, non-derivative sense is the agent's having a conscious perspective. When an agent lacks a perspective, we have no choice but to see it as a system of integrated mechanisms. Of course, the agent's various parts work together to produce movements that aid the organism's survival, as with the sunflower's twisting round to face the sun. But neither these individual mechanisms, nor the being whose behaviour is the result of their work, can possess the kind of goal I have just been describing. Without perspective, we can see movements as purposive in biological terms, but there is nothing overall to which we can attribute a goal that is not derived from natural selection.

As we have seen, when an agent possesses phenomenal consciousness, this confers a personal, or animal, level perspective which allows us to see the agent, rather than parts of the agent, taking in information and interacting with the world. This unifying role played by the perspective ensures that when there is a conscious perspective on the scene, there is more than the activity of sub-animal mechanisms: there is an agent that can be the subject of abilities and goals. It is the perspective that allows us to pick out the kinds of beings that can possess goals.

Let us return to Tinbergen's description of stickleback mating behaviour, to see what a difference the attribution of goals and a perspective makes. Tinbergen's original description of the stickleback's behaviour, in which he describes behaviours being 'released' by certain stimuli, is compatible with sticklebacks lacking a

perspective and possessing goals only derivatively. We can see how the fish's reflexive reactions have a purpose in terms of being naturally selected for survival value, but there is no need to see the fish itself as possessing any goals.

To see the stickleback as exercising abilities when it comes to its mating behaviour, we must see it as having goals and a conscious perspective.[20] If a conscious perspective is on the scene, it makes sense to say that, instead of getting a 'visual stimulus' from the female that induces a reaction, the stickleback *sees* the female and acts in response to what it sees. It now makes sense to see the stickleback as having a goal of mating with the female, and its acting in pursuit of that goal. As well as sub-personal mechanisms with a certain selectional history doing their work, we have a being that is present in the here and now, taking in information about its environment which informs its behaviour. The difference a perspective makes is that there is an animal-level 'centre' unifying the sub-personal mechanisms at work within the stickleback. Instead of the stickleback's sub-animal systems responding to stimuli being the main actors on the scene, the stickleback as a whole agent is present as well.

Of course, the fact that animals (human and non-human) are biological beings means that there is an extent to which all that we do can be traced back to natural selection. I am not claiming that when an agent has a perspective, this means that their goals and behaviour come completely free from evolutionary history. Actions such as my searching the kitchen cupboards because I am hungry can still be linked to the role of hunger in survival. But even though there remains a link to evolutionary history, this case is very different from my pulling back my hand from a hot pan.

The same goes in the case of non-human animals—a dog following its owner around because it wants the owner to throw a ball may be doing so because it has a drive to stick to members of a pack. But this is very different to a case of the dog's panting or wagging its tail—these behaviours are 'derivatively' goal-directed, not done in pursuit of goals of the dog.

20 My claim at this point is not that the stickleback does possess a conscious perspective—my claim is that it is only possible to see the stickleback as possessing abilities if we see it as possessing a conscious perspective.

Having established that possession of a conscious perspective is an indicator of the possession of abilities, we are now a step closer to being able to say where a line should be drawn between those beings that possess abilities and those that do not. The final step before we can answer this question is to say something about how we decide whether a being has a conscious perspective.

3.

Evidence for the possession of a conscious perspective

3.1 Which animals have a conscious perspective?

We began investigating which characteristics are necessary for the possession of abilities in order that the question of which animals have abilities may be answered more easily: we can answer the question by looking for these characteristics in animals and taking them as indicators of the possession of abilities. However, now we have seen that a conscious perspective is required for the possession of ability, it may seem that this method has not made our task any easier. Questions of which animals are conscious are notoriously difficult to answer, and may seem just as—if not more—difficult than the question about abilities that we were asking in the first place.

Deciding which animals possess a conscious perspective requires us getting 'into the heads' of animals. This is a difficult task. Of course, in our everyday contact with animals we ascribe to them experiences and subjectivity—the deer is feeling pain, the mouse smells food. But many argue that we cannot rely on the kind of judgement that is based on generalizing our experiences to non-human animals. For example, Brian Key writes: 'because of the ubiquity of their subjective experiences, humans seem to readily accept the anthropomorphic extension of these mental states to other animals'.[21]

21 (Key 2015) p. 149

There is an attitude to investigating consciousness, often accepted by those who make accusations of anthropomorphism, that makes the question of which beings are phenomenally conscious and which are not appear incredibly difficult, if not impossible, to answer. But if we take an alternative approach to thinking about consciousness we will see that the question can be answered with relative simplicity. Or, at least, that there are methodologies that will lead to the question being answered, even if we have no answer at the present time. Questions about which beings are conscious are not intractable.

There are two ways of thinking about consciousness that we must avoid if our investigation into which animals have a conscious perspective is to be successful. One of these suggests that we must abandon 'anthropomorphic' descriptions of animal behaviour in favour of more 'neutral' language if we are to accurately answer questions about animals' inner lives. The other problematic thought is that phenomenal consciousness has no function and plays no role in behaviour: it is epiphenomenal. The thought here is that animals could behave in exactly the same way whether or not they possess consciousness, just as David Chalmers' philosophical zombies behave indistinguishably from humans with consciousness.[22] Once we note these ways of thinking about consciousness and are careful to avoid the pitfalls they may lead us to, we can see that deciding which animals have conscious perspectives is not such a difficult task after all.

Eileen Crist notes the distorting influence that language can have on our picture of animal agency.[23] She argues that language plays a central role in our thought about animals, writing that 'the powerful impact of language is not [...] manifest only at the level of abstract ideas, but also at that of the reader's perceptual and affective experience.'[24] One of Crist's central aims is to challenge the ideas that using the same language to describe animal agents as we use to describe human agents is a mistake, and that describing animals' behaviour using 'neutral' scientific terms is more correct. She argues, rather, that each way of describing animals carries with it its own picture of the conduct of animals, and that no description is closer to an objective account of animal behaviour—something that does not exist.

22 (Chalmers 1996)

23 (Crist 1991)

24 Ibid. p. 208

We have already seen the effects language can have on our perception of the conduct of animals, through comparing Warington's and Tinbergen's descriptions of stickleback mating behaviour earlier in this chapter. There, technical terms like 'visual stimulus', 'stimulate', 'release' and 'induce' jarred with Warington's description of the male fish 'endeavor[ing] to push' the female toward the nest, and presented a picture that seemed removed from a common-sense perception of the scene.[25] But in this case, Tinbergen's use of technical language is the result of his attempting to present what he would accept is a decidedly non-neutral picture of animal conduct, on which the fishes' behaviour is governed by stimulus-response mechanisms. It is no surprise, then, that his language produces a certain picture of the sticklebacks' agency.

A better example of Crist's point can be found in Konrad Lorenz's description of a starling that has been kept indoors and has developed a habit of behaving as if it is hunting, even though there is nothing for it to hunt. Lorenz is careful to keep his description neutral, but despite this his language conveys a particular picture of the starling's conduct and mental life:

> *With head and eyes the bird made a motion as though following a flying insect with its gaze; its posture tautened; it took off, snapped, returned to its perch, and with its bill performed the sideways lashing, tossing motions with which many insectivorous birds slay their prey against whatever they happen to be sitting upon. Then the starling swallowed several times, whereupon its closely laid plumage loosened up somewhat, and there often ensued a quivering reflex, exactly as it does after real satiation.*[26]

Lorenz uses little technical terminology, giving precise and vivid descriptions of the starling's movements. It may seem that by describing only the starling's movements, and avoiding any comment on the less certain issue of what is going on 'inside its head', he is providing us a neutral account of its behaviour from which we can draw our own conclusions. However, even this description carries with it its own way of influencing our perception of the starling's conduct. By avoiding talk of the starling's inner life and focusing only on its behaviour, Lorenz presents a picture more in tune with the thought that the bird is simply manifesting a disposition. Although not

25 (Warington 1852) p. 277; (Tinbergen 1951) p. 47

26 (Lorenz 1957) p. 143

ruling out that the starling has an inner life and is exercising its abilities, it makes the starling's conduct seem like poor evidence for this.[27]

The second thought we must be careful to avoid is that phenomenal consciousness plays no role in behaviour. For the philosophers who hold this view—those, for example who accept that Chalmers' philosophical zombies are conceivable—no animal behaviour can be used as evidence for the possession or lack of phenomenal consciousness. The thought is that since animals' behaviour and physiology would be exactly the same no matter whether or not the animal had phenomenal consciousness, nothing that animals do and no features of their physiology can be used in deciding whether animals have consciousness. On this way of thinking about consciousness, it seems that the question of which animals possess a conscious perspective is intractable. Donald Griffin and Gayle Speck complain that 'many behavioural scientists have been extremely reluctant to consider non-human consciousness on the grounds that it is impossible to obtain objective evidence about subjective experience.'[28]

Our question about animal consciousness is predicated on the idea that phenomenal consciousness does have a functional role—and so we cannot entertain a view like Chalmers'. The reason we are trying to determine which animals have a conscious perspective is that having phenomenal consciousness makes a difference in agents' behaviour—it makes the possession of abilities possible. It is also worth noting that much scientific investigation into animal minds is also built on the thought that consciousness does have a role in animals' behaviour. We must either reject the view that consciousness is epiphenomenal or simply halt our investigations here.

With these two points in mind, we can see that we can have a fruitful debate about consciousness. This is illustrated by looking at an example of a debate that shows how we can argue about consciousness, provided we accept it has a functional role and are careful to note the effect the language we use to describe animals' conduct

27 Primatologist Frans de Waal advances a complaint similar to Crist's. In de Waal's case this is against those who make an effort to employ 'neutral' terminology in descriptions of ape behaviour: *Dubbing an ape's kiss "mouth to mouth contact" so as to avoid anthropomorphism deliberately obfuscates the meaning of the behaviour. It would be like assigning Earth's gravity a different name from the moon's, just because we think Earth is special. Unjustified linguistic barriers fragment the unity with which nature presents us.* (de Waal 2016, p. 21)

28 (Griffin & Speck 2004) p. 5

has on our perception of their inner lives. When we see that phenomenal consciousness is not epiphenomenal, we can use a wide range of experimental and observational evidence to decide questions about which animals possess conscious perspectives. And when we are aware of the perspectival effects that 'scientific' descriptions of animals' behaviour can have, we can ensure that they do not unduly prejudice our perception of animals' conduct.

3.2 What counts as evidence of a conscious perspective?

We are nearly ready to assess the available evidence to decide which animals possess a conscious perspective and thereby possess abilities. The discussion in the previous section has shown us that, far from seeing a conscious perspective as a mysterious epiphenomenon, we should see it as something that plays a role in animal behaviour. Therefore the evidence we will assess in deciding which animals possess conscious perspectives will come from observations of animals' behaviour.

The question now is what kind of behaviour is indicative of the possession of a conscious perspective—what behaviour should we look for in drawing our conclusions about consciousness? Michael Tye proposes a test to decide whether an animal is conscious based on its behaviour. A conscious perspective allows for flexibility in behaviour, so where animals display flexible behaviour—different responses to the same stimuli—this is evidence they possess a conscious perspective.[29] Tye's thought is that possessing an inner life allows for more varied reactions to the same stimuli. He argues that, conversely, a rigidity in reaction to stimuli indicates a lack of an inner life—a lack of conscious perspective.

Tye's claim that flexible behaviour is an indicator of the possession of a conscious perspective seems broadly correct, and fits in with my claim in this chapter that a conscious perspective is necessary for the possession of abilities. If conscious perspective is what makes the possession of abilities—and the flexible behaviour that is abilities' exercise—possible, then it makes sense that flexible behaviour by an animal is an indicator that the animal possesses a conscious perspective. Tye's test

29 (Tye 2016) p. 162

for the possession of a conscious perspective is in accord with the idea that ability and conscious perspective are linked.

Even so, we may seem to be going round in circles. We set out to make the task of deciding which animals possess abilities an easier one by looking at characteristics that might indicate the possession of abilities. The aim of looking at which animals possess a conscious perspective is to enable us to more easily decide which animals possess abilities. Now we are looking for a certain kind of flexible behaviour, the hallmark of a being's possessing abilities, in order to decide whether an animal possesses a conscious perspective. What is the use of looking for conscious perspective in order to make our investigation into the possession of abilities easier, if to decide whether an animal possesses a conscious perspective we must look at the kind of behaviour that indicates the possession of an ability?

Despite this appearance of circularity, the approach of looking for evidence of a conscious perspective as a way of determining which animals possess abilities *does* aid our investigation. It is true that flexible behaviour is a key indicator that an animal possesses a conscious perspective, and that drawing conclusions about whether an animal possesses a conscious perspective will involve looking for this behaviour. But asking whether an animal has a conscious perspective opens up far more lines of investigation than would be available if we were to directly look at whether it possesses abilities. There is a wealth of literature assessing empirical evidence for animals' possession of consciousness, which includes not just observations of behaviour but observations of animals' physiology and brain structures. If we were looking directly at ability these avenues would not be open to us.

So, we should look at flexible behaviour as evidence for the possession of a conscious perspective—but we should also consider other available evidence from experimentation and investigation into the animal's physiology. With these multiple sources of evidence we have more to draw on when deciding whether an animal possesses a conscious perspective—evidence that can be drawn on to ultimately determine whether an animal possesses abilities.

Consider the following example.[30] When fish of certain species come into contact with a harmful stimulus they will behave in a certain way, removing the affected body part and moving away from the stimulus. In the case of some harmful stimuli certain species will engage in other behaviour as well: an experiment carried out by Lynne Sneddon involved injecting bee venom or acid into the lips of trout, who would then rub their lips against the sides and base of their tanks.[31] In a further experiment carried out by Sneddon, trout were captured when they approached a food ring in response to a light cue in their tank. Some of the trout were injected with harmless saline, some with acetic acid, some with bee venom, and some were not injected at all. After they had been put back in the tank, the trout injected with acid or venom took the longest time before they started feeding in response to the light cue again.

What conclusions can we draw from this about the possibility that trout possess a conscious perspective? When fish move away from conscious stimuli this can seem similar to a human's drawing her hand away from a hot pan after accidentally touching it. In the human case this kind of behaviour usually indicates that the person is feeling pain. So, we might take the fish's behaviour as evidence that the fish is feeling pain. However, as we have seen, simple reflex movements are not good indicators that a conscious experience is occurring. Experimental evidence further supports the idea that we should not take this behaviour as evidence that the fish is undergoing a conscious experience.

Some experiments on fish involve removing the cerebral hemispheres—the parts of the brain that would contain the structures responsible for the sensation of pain if it were present.[32] Even when these parts of the brain are removed, the fish continue to behave in exactly the same way as they did before in response to the harmful stimulus. Since the behaviour is the same in both cases, this seems to show that the fish's behaviour of moving away from harmful stimuli cannot be taken as evidence that the fish is feeling pain. Pain seemingly plays no role in this behaviour.

---

30 The examples that follow focus ask whether animals have individual experiences, like the experience of pain, rather than whether these animals possess a conscious perspective. As we saw in Section 2.1, having an individual experience such as the experience of pain is sufficient for possessing a conscious perspective.

31 (Sneddon 2013)

32 (Rose 2002). Brian Key (2015, p. 152) surveys a number of similar experiments in which fish continue to behave in the same way following decerebration.

However, looking at the second set of behaviours displayed in Sneddon's injection experiment, we do seem to have evidence in favour of the conclusion that trout possess a conscious perspective. In avoiding their food after the injection, the trout appear to behave flexibly—responding in different ways to the same stimuli. This indicates that they possess a conscious perspective from which they associate the sensation of pain with the food ring. And, whereas the behaviour of moving away from harmful stimuli is still carried out when the frontal hemispheres are removed, there is no evidence that the behaviour of avoiding the food ring after being injected with acid or venom continues.[33] This suggests the fish's conscious experiences are playing a role in their behaviour.

The above discussion shows how we can draw conclusions about the conscious perspective, and thereby the abilities, of fish by looking at a range of behavioural and experimental evidence. There is a wealth of evidence available that can help us draw conclusions about the possession or lack of conscious perspective in non-human animals.

It is important at this point to make clear that we cannot expect to find conclusive evidence of an animal's possession or lack of a conscious perspective: if we set out to find the piece of evidence that allows us to state with certainty that an animal possesses a conscious perspective, we are bound to end up disappointed. Although, as we have seen, an animal's possession or lack of a conscious inner life is revealed in its behaviour and seems determined by its physiology, consciousness is still in some sense a private phenomenon. We cannot directly determine what another being is experiencing. As demonstrated Lorenz and Tinbergen who avoided 'anthropomorphic' descriptions of animal behaviour, if one looks hard enough there is often an alternative explanation of animal behaviour involving drives, instincts or dispositions.

So, instead of looking for behavioural or physiological evidence of an animal's possessing a conscious perspective that will somehow blow all other explanations out of the water, we should instead ask whether the evidence makes it more or less likely that an explanation in terms of the animal's possessing a conscious perspective is the correct one. This is the approach taken by Donald Griffin and Gayle Speck,

33 (Tye 2016) p. 104

who investigate animal consciousness in terms of the 'probability of awareness (pA)': 'If we have complete certainty that a given animal has a particular conscious experience, then pA=1.0, and pA=0 means that we know with certainty that it does not.'[34] We need not numerically quantify the likelihood of an animal's possessing a conscious perspective as Griffin and Speck do, but should simply bear in mind that we are looking at whether an animal's possessing a conscious perspective is the best explanation given the available evidence—not whether the evidence rules out all other possibilities except that the animal possesses a conscious perspective.

We are now in a position to apply our approach in investigating the possession or lack of a conscious perspective in a variety of animals. Looking at the available empirical evidence will show us roughly where we should draw the line between those animals that possess conscious perspectives and abilities, and those that do not.

3.3 Empirical investigation into animal consciousness

Let us begin at the easier end of the scale. It will not seem too much of a stretch to suppose that all mammals possess conscious perspectives and therefore possess abilities. Evidence both from the behaviour and the physiology of mammals points us towards this conclusion.

The flexible behaviour of many mammals, such as problem-solving and social interaction, is familiar to us, whether in bonobos or cats. To give a concrete example of the kind of behaviour that might evidence the possession of a conscious perspective, let us look at a less familiar case: the digging behaviour of naked mole rats.

Naked mole rats search for plants and tubers in the ground by digging tunnels with their large front teeth. When digging tunnels in a material that creates fine dust, they have been observed placing a piece of wood or plant behind their front teeth and in front of their lips, thereby creating a barrier that prevents them inhaling any of the

34 (Griffin & Speck 2004) p. 6

dust created by their tunnelling.[35] Naked mole rats do not use such a barrier when they are digging in a material that does not create fine particles that could be inhaled. If their barrier breaks or is knocked out of place, they will readjust it, search for something else to use as a barrier, or stop digging altogether. In reporting observations of this behaviour, Gabriela Shuster and P. W. Sherman argue that this tool use, which involves problem-solving skills, could not be achieved by the mole rat without possessing perceptual consciousness.[36]

Although this behaviour makes it more probable that the mole-rat possesses a conscious perspective, one may attempt to explain it in terms of innate drives and manifestations of dispositions. But the similarity of humans' and non-human mammals' physiology takes us closer toward the conclusion that naked mole-rats and other mammals possess conscious perspectives.

The thalamus and cortex in human brains are thought to be essential for consciousness.[37] One main reason these structures are thought to be linked to consciousness is thata lesion in some regions of this 'thalamocortical system' results in loss of specific perceptual knowledge—for example, it may leave the subject unable to perceive faces.[38] Bilateral lesions in the thalamocortical system lead to the loss of consciousness altogether. Like humans, all mammals have a highly developed thalamocortical system, suggesting that they too have conscious perspectives. The electrical activity in the brains of humans and non-human mammals is also strikingly similar, with an identical contrast between the activity seen during waking and sleeping states—to the extent that the findings of studies of the electrical activity in the brains of mammals are often applied to the case of humans.[39] Added to observation of their flexible behaviour, this encourages us to conclude that all mammals possess conscious perspectives.

Now let us turn to birds, whose physiology and behaviour seems in some ways more removed from that of humans and other mammals. Birds lack a thalamocortical system, so in this respect are quite different from mammals. However, some systems

35 (Shuster & Sherman 1998)

36 Ibid.

37 (Edelman & Tononi 2000) p. 54

38 Ibid.

39 Ibid.

in the brain of birds are taken to be homologous to the thalamocortical system in mammals. Two brain structures in birds, the Wulst and the dorsal ventricular ridge, have very similar neurochemical traits and circuitry to the thalamocortical system. Lesions to these areas impair behaviours that require more complex cognitive capacities, such as memory dependent tasks.[40] Further to this, the electrical patterns measured during waking states in birds are similar to those in mammal brains—although the pattern measured in birds during sleeping states is very different to that seen in sleeping mammals.[41]

When we add this evidence to the well-documented cases of tool use by crows—which can solve a range of 'puzzles' involving getting food from hard to reach places using sticks—and social interaction by parrots, the conclusion that these birds possess conscious perspectives becomes more likely.[42] The physiological evidence is not as clear as in the case of mammals, however, and so we should look at the behaviour of different species of bird on a more case-by-case basis.

Now we move onto more difficult cases—to the realm of insects. We may have been fairly inclined to attribute abilities to mammals and birds, but our intuitions about insects are less clear—we may waver between thinking of insects as agents in control of their bodies, or complex bundles of biological stimulus-response mechanisms. In answering the question of whether insects possess conscious perspectives, and abilities, empirical evidence becomes even more valuable.

One experiment into the behaviour of fruit flies appears to suggest that they possess conscious perspectives. When these flies are given some forms of analgesic, the effects seem similar to those on humans and other mammals.[43] The flies are placed in a test-tube in which there is a light gradient. In normal conditions, they will move towards the area with stronger light. But if the centre of the tube is heated, the flies will not pass through the heated area. When an analgesic is administered to the flies, they move through the heated part of the tube towards the light. This could be

40 (Butler et al. 2005) p. 932-933

41 (Beshkar 2008) p. 22

42 Such cases are surveyed in (Beshkar 2008) pp. 20-21

43 (Dimitrijevic et al. 2005)

interpreted as the analgesic numbing the flies' sensation of pain, allowing them to pass through the heated section of the tube unhindered.

This experiment alone seems weak evidence of the possession of a conscious perspective. There is no flexible behaviour on the part of the flies—there are only different reactions to the heat stimulus depending on different conditions. Further, it may be the case that the analgesic, rather than numbing pain that the flies are experiencing, simply changes the make-up of the flies' physiology in some other way that makes them more inclined to ignore harmful stimuli. For example, the analgesic may simply make the flies 'confused'.

Other, more promising, evidence, is discussed by Michael Tye.[44] He presents a study apparently showing that fruit flies can learn to associate an odour with an electric shock. After being exposed to a particular odour followed by an electric shock several times, the fly will then avoid that odour for up to twenty-four hours. One way to interpret this is that the flies are experiencing pain in response to the electric shock and associating it with the experience of the odour. This apparent association is a case of a more complex behaviour than that displayed in the heated test tube study discussed above, and seems better evidence for the fly's possessing a conscious perspective. We are still, however, faced with the fact that the physiology of fruit flies is far removed from that of humans and other mammals—which lessens the probability that fruit flies possess a conscious perspective and suggests we should look for other explanations for this behaviour in terms of dispositions.

Some recent research may provide evidence that flies' physiology is such that it can support a conscious perspective. The gene pathways that are responsible for the avoidance of harmful stimuli in flies are similar to those that set pain thresholds in mice—and even in humans. It has been found that deletion of these genes in both flies and mice resulted in reduced reaction to harmful heat stimuli.[45] Variations of these genes in humans were also found to be linked to different levels of pain sensitivity in tests. The researchers carrying out the study conclude that 'these data reinforce the extraordinary conservation of the neurobiological mechanisms of nociception, from its manifestation as avoidance of damage in primitive creatures

44 (Tye 2016) p. 140

45 (Neely et al. 2010)

like flies to the complex sensation of pain in humans."[46] Of course, the possession of a conscious perspective by fruit flies still does not seem a very likely prospect. But these findings combined with observations of flies' behaviour suggest we should entertain and further investigate the idea that fruit flies possess at least a limited range of abilities.

An example where it seems the probability of an animal possessing a conscious perspective is very low comes in the case of the flatworm. In fact, these worms do display behaviour that we may interpret as involving more than a simple stimulus-response mechanism. In one experiment, worms are placed in a glass tube and given electrical shocks as they travel through it.[47] This process is carried out several times, and in successive trials the worms will crawl through the tube more and more slowly, until eventually they will not move through the tube at all. This behaviour will be remembered for up to six hours. As in the case of the fruit flies' association between electric shock and odour, the flatworms' avoidance behaviour—and the fact that it is remembered—is suggestive of the worms' having pain experiences that they associate with crawling through the glass.

However, what speaks against this interpretation is what we know about the physiology of flatworms: that organisms like this have only around 300 neurons, compared to a fruit fly's 250,000, a brown rat's 2 million, or a human's 86 million.[48] This stark neurophysiological divide between flatworms and those animals to which the attribution of consciousness appears more likely serves to lower the probability that flatworms possess a conscious perspective. We should therefore favour other explanations of the worms' avoidance behaviour, in terms of stimulus-response mechanisms that manifest dispositions rather than in terms of the worms' possessing abilities.

46 (Neely et al. 2010) p. 635

47 (Clark 1965)

48 (Chatterjee & Sinha 2008) p. 146

Conclusion

Looking at the possession of a conscious perspective as an indicator of the possession of abilities has enabled us to say roughly where the lines are drawn between those animals that possess abilities and those that merely manifest dispositions. We have seen that the line is drawn very far from human agents, and that even certain insects may be agents exercising abilities, in control of their bodies and possessing a conscious perspective.

This investigation has made clear that we cannot rely solely on our observations of an animal's movements to determine whether an animal is exercising an ability or manifesting a disposition—and, moreover, we cannot rely solely on isolated observations of particular behaviours. From a neutral perspective it looks the same to us whether a fly is exercising its abilities in trying to avoid a painful electric shock, or manifesting a disposition to avoid harmful stimuli.[49] We must put together all the evidence we have—from observing as much behaviour as possible, as well as drawing on what we know about the animal's brain and physiology—in order to determine what is actually happening—whether the animal is exercising abilities or manifesting dispositions.

There are some animals whose conduct we cannot help but liken to ours. There are animals we spend so much time with, such as cats and dogs, and animals whose behaviour seems so similar to ours, such as chimpanzees, that our understanding of their behaviour is suffused with comparisons to our own exercises of ability. In other cases, as with fruit flies, we have no strong intuitions, or may oscillate between seeing an animal as an agent in control of its body and a bundle of stimulus-response mechanisms. Answering the question of whether these animals possess abilities shows us how we should view their conduct. If the evidence for flies possessing a conscious perspective begins to stack up, we learn something new about what is really happening in cases like the heated test-tube experiment. We learn that the fly is trying to avoid pain or harm: it is exercising its abilities to move away from the heat. An affirmative answer to the question of whether flies have abilities would show us that despite the various differences between the agency of flies and humans, they,

49 Although, as Crist (1991, p. 208) contends, it is unlikely that we can ever adopt a truly neutral perspective in describing animal behaviour.

like us, are agents with perspectives on the world, interacting with their environments through exercising bodily abilities.

Conclusion

The foundational idea of this thesis has been that, by looking at what human agency has in common with the agency of non-human animals and inanimate objects, we can gain an interesting new perspective on the powers essential to human agency.

We have found that central to human agency are bodily abilities of a kind also possessed by many non-human animals. What distinguishes these powers from dispositions is not that abilities' exercise is never necessitated by certain conditions—we have seen that dispositions, too, are never necessitated to manifest in any conditions. Instead, it is the fact that ability involves self-movement—an embodied agent moving its own body—that distinguishes these powers from dispositions. Possessing these bodily abilities does not require high-level mental capacities. Rather, it requires only a conscious perspective, leading to the conclusion that animals very different to humans may possess these powers.

Much of the picture developed over the course of this thesis has been pieced together from accounts presented by other philosophers. This is an opportunity to retrace the path taken through these accounts to arrive at a view of human agency and its place among agency of other kinds.

Mumford and Anjum's account of dispositions showed us why the difference between dispositions and abilities, and in consequence the difference between human agency and the agency of inanimate objects, does not lie in necessity. They argue that the modality of dispositions is stronger than logical possibility but weaker than necessity, since dispositions can always be prevented from manifesting. Drawing on this argument allowed us to see that our concept of disposition is not the concept of a power necessitated to manifest in certain conditions, even if we acknowledge that, in reality, dispositions are necessitated to manifest by a totality of conditions at a time. I have suggested that we conceive of dispositions, like abilities, as powers whose manifestations are not necessitated.

A discussion by Markus Schrenk served to highlight how the view that dispositions 'push' or 'tend' towards their manifestations can introduce the idea of a fundamental

difference between abilities and dispositions. In light of this view, it appeared that dispositions have three possible states—not manifesting, pushing towards a manifestation and manifesting—whilst abilities have only two—not being exercised and being exercised. Showing it is not the case that dispositions have three states whilst abilities have only two involved developing an account of multi-track dispositions. On this account, the manifestations of dispositions are effects, and any disposition may have different manifestations with different manifestation partners. This account enabled us to see how it could be that a disposition's pushing *is* its manifestation: a process that could be altered or prevented from reaching its end point through the activity of other powers.

Anton Ford's introduction of the notion of transaction provided the opportunity to say more about self-movement, a key feature of the exercises of abilities that separates them from the manifestations of dispositions. We accepted Ford's view that human agents transact with other objects, but saw that Ford's downplaying of self-movement led to a picture of human agency that contained a gap: it left us with no way to articulate the embodied nature of human agency. To remedy this, a multi-track picture of abilities was developed along the same lines as the multi-track view of dispositions. This allowed us to see that all transactions are exercises of bodily abilities, thus demonstrating that there need be no inconsistency in accepting both that human agents transact and that human agency is fundamentally embodied.

Looking at views of animal agency in the scientific literature allowed us to say more about which characteristics of an animal are required in order for it to possess abilities. We saw that possessing abilities requires only a conscious perspective, rather than the higher-level mental capacities distinctive of human agency. When there is a conscious perspective on the scene, there is more than the activity of sub-animal mechanisms: there is an agent that can be the subject of abilities and goals.

A conception of the bodily nature of our agency may be overlooked when all that is considered is what separates human agency from agency of other kinds. We are embodied beings, and this is reflected in the way we bring about change in the world through our bodily abilities. Although there are many ways in which the things that we do are predicated on mental capacities unique to human agents, our agency is grounded in powers of a kind that human and non-human agents have in common.

Bibliography

Aguilar, J. H. & Buckareff, A. A. 2010. 'The Causal Theory of Action: Origins and Issues' in Aguilar, J. H. & Buckareff, A. A. (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action*. MIT Press.

Allen, A. & Beckoff, M. 1999. *Species of Mind: Philosophy and Biology of Cognitive Ethology*. MIT Press.

Alvarez, M. 2013. 'Agency and Two-Way Powers', *Proceedings of the Aristotelian Society* 113(1): 101-121

Alvarez, M. & Hyman, J. 1998. 'Agents and their Actions', *Philosophy* 73(2): 219-245

Anscombe, G. E. M. 1957. *Intention*. Blackwell.

Baars, B. J. 2005. 'Subjective experience is probably not limited to humans: The evidence from neurobiology and behaviour', *Consciousness and Cognition* 14(1): 7-21

Beshkar, M. 2008. 'Animal Consciousness', *Journal of Consciousness Studies* 15(3): 5-33

Bird, A. 1998. 'Dispositions and Antidotes', *The Philosophical Quarterly* 48(191): 227-234

Bird, A. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford University Press.

Bishop, J. 1986. 'Is Agent-Causality a Conceptual Primitive?', *Synthese* 67(2): 225-247

Bishop, J. 1989. *Natural Agency*. Cambridge University Press.

Bishop, J. 2010. 'Scepticism about Natural Agency and the Causal Theory of Action' in Aguilar, J. H. & Buckareff, A. A. (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action*. MIT Press.

Brand, M. 1984. *Intending and Acting: Towards a Naturalized Action Theory*. MIT Press.

Bratman, M. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press.

Butler, A. B., Manger, P. R., Lindahl, B. I. B. & Arhem, P. 2005. 'Evolution of the neural basis of consciousness: a bird-mammal comparison', *Bioessays*, 27(9): 923-936

Carruthers, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Oxford University Press.

Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press.

Chalmers, D. 1996. *The Conscious Mind*. Oxford University Press.

Chatterjee, N. & Sinha, S. 2008. 'Understanding the mind of a worm: hierarchical network structure underlying nervous system function in C. elegans' in Banerjee, R. & Chakrabarti, B. K. (eds.), *Models of Brain and Mind: Physical, Computational and Psychological Approaches*. Elsevier.

Chisholm, R. 1966. 'Freedom and Action', Lehrer, K. (ed.), *Freedom and Determinism*. Random House.

Clark, R. 1965. 'The Learning Abilities of Nereid Polychaetes and the Role of the Supra-Oesophagenal Ganglion.' *Animal Behavior* 13(Supplement 1): 89-100

Crist, E. 1991. *Images of Animals*. Temple University Press.

Danto, A. 1976. 'Action, Knowledge and Representation' in Brand, M. & Walton, D. (eds.), *Action Theory: Proceedings of the Winnipeg Conference on Human Action, held at Winnipeg, Manitoba, Canada, 9-11 May 1975*. D. Reidel Publishing Company

Davidson, D. 1980. 'Agency', in Davidson, D. *Essays on Actions and Events*. Oxford University Press.

Davidson, D. 2004. *Problems of Rationality*. Oxford University Press.

de Waal, F. 2016. *Are We Smart Enough to Know How Smart Animals Are?* Granta.

Di Nucci, E. 2013. 'Action, Deviance and Guidance', *Abstracta*, 7(2): 41-59

Dimitrijevic, N., Dzitoyeva, S., Satta, R., Imbesi, M., Yildiz, S., & Manev, H. 2005. 'Drosophila GABA(B) Receptors Are Involved in Behavioral Effects of Gamma-Hydroxybutyric Acid (GHB)', *European Journal of Pharmacology* 519(3): 246-252

Douskos, C. 2013. 'Settling and Bodily Control', *Inquiry* 56(6): 639-652

Dretske, F. 1988. *Explaining Behaviour: Reasons in a World of Causes*. MIT Press.

Dyke, H. 2007. *Metaphysics and the Representational Fallacy*. Routledge.

Eddington, A. S. 1928. *The Nature of the Physical World*. Macmillan

Edelman, G. M. 2003. 'Naturalizing consciousness: A theoretical framework', *Proceedings of the National Academy of Sciences of the United States of America*, 100(9): 5520-5524

Edelman, G. M. & Tononi, G. 2001. *A Universe of Consciousness*. Basic Books

Fara, M. 2005. 'Dispositions and Habituals', *Nous*, 39(1): 43-82

Fara, M. 2008. 'Masked Abilities and Compatibilism', *Mind*, 117(468): 843-865

Frankfurt, H. 1978. 'The Problem of Action', *American Philosophical Quarterly*, 15(2): 157-162

Ford, A. 2013. 'Is Agency a Power of Self-Movement?', *Inquiry*, 56(6): 597-610

Ford, A. 2014. 'Action and Passion', *Philosophical Topics,* 42(1): 13-24

Ford, A. 2016. 'The Province of Human Agency', *Nous*, online early view

Griffin, D. R. & Speck, G. B. 2004. 'New evidence of animal consciousness', *Animal Cognition* 7: 5-18

Haddock, A. 2005. 'At one with our actions, but at two with our bodies', *Philosophical Explorations*, 8(2): 157-172

Heil, J. 2003. *From an Ontological Point of View*. Oxford University Press.

Heil, J. 2005. 'Real Tables', *The Monist* 88(4): 493-509

Heil, J. 2012. *The Universe As We Find It*. Oxford University Press.

Hornsby, J. 2009. 'Physicalism, Conceptual Analysis and Acts of Faith' in Ian Ravenscroft (ed.), *Minds, Ethics and Conditionals: Themes from the Philosophy of Frank Jackson*. Oxford University Press.

Hornsby, J. 2010. 'Trying to Act' in O'Connor, T. & Sandis, C. (eds.), *A Companion to the Philosophy of Action*. Blackwell.

Hornsby, J. 2011. 'Actions in Their Circumstances' in Ford, A., Hornsby, J. & Stoutland, F. (eds.), *Essays on Anscombe's Intention.* Harvard University Press.

Kenny, A. J. P. 1989. *The Metaphysics of Mind.* Oxford University Press.

Key, B. 2015. 'Fish do not feel pain and its implications for understanding phenomenal consciousness', *Biology & Philosophy* 30(2): 149-165

Lavin, D. 2013. 'Must There Be Basic Action?', *Nous*, 47(2): 273-301

Lehrer, K. 1968. 'Cans without Ifs', *Analysis*, 29(1): 29-32

Lewis, D. 1997. 'Finkish Dispositions', *The Philosophical Quarterly* 47(187): 143-158

Levy, N. 2013. 'Intentional Action First', *Australasian Journal of Philosophy*, 91(4): 705-718

Locke, D. 1973. "Natural Powers and Human Abilities", *Proceedings of the Aristotelian Society* 74: 171-187

Lorenz, K. 1957. 'The nature of instinct' in Schiller, C. H. and Lashley, K. S. (eds.), *Instinctive behavior.* International University Press.

Lowe, E. J. 2010. 'On the Individuation of Powers' in Anna Marmodoro (ed.), *The Metaphysics of Powers: Their Grounding and their Manifestations*. Taylor & Francis.

Maier, J. 2014. 'Abilities', in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2014 Edition*) URL = https://plato.stanford.edu/archives/fall2014/entries/abilities/

Martin, C. B. 1994. 'Dispositions and Conditionals', *The Philosophical Quarterly* 44(174): 1-8

Mayr, E. 2011. *Understanding Human Agency*. Oxford University Press.

McCann, H. J. 1974. 'Volition and basic action' in McCann, H. J. (ed.), *The works of agency: On human action, will and freedom*. Cornell University Press.

McDowell, J. 'The content of perceptual experience', *The Philosophical Quarterly*, 44(175): 190-205

McKitrick, J. 2010. 'Manifestations as Effects' in Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and their Manifestations*. Taylor & Francis.

McKitrick, J. 2013. 'Book Symposium: Causes as powers', *Metascience*, 22(3): 545-559

Melden, A. I. 1961. *Free Action*. Routledge & Kegan Paul.

Mele, A. R. & Moser, P. K. 1994. 'Intentional Action', *Nous* 28(1): 39-68

Molnar, G. 2003. *Powers: a study in metaphysics*. Oxford University Press.

Mumford, S. & Anjum, R. L. 2011a. *Getting Causes From Powers*. Oxford University Press.

Mumford, S. & Anjum, R. L. 2011b. 'Dispositional Modality' in Gethmann, C. F. (ed.), *Lebenswelt und Wissenschaft, Deutsches Jahrbuch Philosophie 2*. Meiner Verlag.

Mumford, S & Anjum, R. L. 2013. 'Author's reply', *Metascience*, 22(3): 545-559

Nagel, T. 1974. 'What is it like to be a bat?', *Philosophical Review* 83(4): 435-450

Neely, G., Hess, M., Costigan, M., et al. 2010. 'A Genome-Wide Drosophila Screen for Heat Nociception Identifies α2δ3 as an Evolutionarily Conserved Pain Gene.' *Cell* 143(4): 628-638

O'Connor, T. 1999. *Persons & Causes: The Metaphysics of Free Will*. Oxford University Press.

O'Shaughnessy, B. 1973. 'Trying (as the mental 'pineal gland')', *Journal of Philosophy*, 70(13):365-386

Pietroski, P. M. 1998. 'Actions, adjuncts, and agency', *Mind* 107(425): 73-111

Quine, W. V. O., 1960. *Word and Object*. MIT Press.

Rose, J. D. 2002. "The Neurobehavioral Nature of Fishes and the Question of Awareness of Pain." *Reviews in Fisheries Sciences* 10(1): 1-38

Ruben, D. 2013. 'Trying in Some Way', *Australasian Journal of Philosophy*, 91(4): 719-733

Russell, B. 1954. *Human Society in Ethics and Politics.* Allen & Unwin.

Ryle, G. 1949. *The Concept of Mind.* Penguin.

Schlosser, M. E. 2011. 'Agency, Ownership and the Standard Theory', in Aguilar, J., Buckareff, A. & Frankish, K. (eds.) *New Waves in the Philosophy of Action.* Palgrave Macmillan.

Schrenk, M. 2009. 'Hic Rhodos, hic salta! How does the Dispositionalist deal with Antidotes?' in Damschen, G., Schnepf, R. & Stüber, K. R. *Debating Dispositions: Issues in Metaphyiscs, Epistemology and Philosophy of Mind.* De Gruyter.

Shuster, G. & Sherman, P. W. 1998. 'Tool use by naked mole-rats', *Animal Cognition* 1(1): 71-74

Small, W. 2016. *Agency and Abilities.* Paper presented at the Royal Institute of Philosophy London Lecture Series.

Small, W. [Forthcoming] 'Bodily Movement and its Significance', *Philosophical Topics*

Smith, A. D. 1988. 'Agency and the essence of actions', *The Philosophical Quarterly*, 38(153): 401-421

Sneddon, L. U. 2003. 'The Evidence for Pain in Fish: The Use of Morphine as an Analgesic.' *Applied Animal Behaviour Science* 83(2): 153-162

Steward, H. 2009. 'Sub-intentional actions and the over-mentalization of agency' in Sandis, C (ed.) *New Essays on the Explanation of Action.* Palgrave Macmillan.

Steward, H. 2012. *A Metaphysics for Freedom.* Oxford University Press.

Steward, H. 2013. 'Responses', *Inquiry* 56(6): 681-706

Stout, R. 1997. 'Processes', *Philosophy* 72(279): 19-27

Stout, R. 2007. 'Two ways to Understand Causality in Agency' in Leist, A. (ed.), *Action in Context.* De Gruyter.

Stoutland, F. 1985. 'Davidson on Intentional Behaviour' in LePore, E. & McLaughlin, B. P. (eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson.* Blackwell.

Stoutland, F. 2011a. 'Anscombe's Intention in Context' in Ford, A., Hornsby, J. & Stoutland, F. (eds.), *Essays on Anscombe's Intention.* Harvard University Press.

Stoutland, F. 2011b. 'Interpreting Davidson on Intentional Action' in Malpas, J. (ed.), *Dialogues with Davidson: Acting, Interpreting, Understanding.* MIT Press.

Taylor, R. 1968. *Action and Purpose.* Humanities Press.

Tinbergen, N. 1951. *The Study of Instinct*. Oxford University Press.

Tye, M. 1995. *Ten Problems of Consciousness: A representational theory of the phenomenal mind*. MIT Press.

Tye, M. 2016. *Tense Bees and Shell-Shocked Crabs*. Oxford University Press.

Van Inwagen, P. 1983. *An Essay on Free Will*. Oxford University Press.

Velleman, D. 1992. 'What Happens When Someone Acts?' *Mind* 101(403): 461-481

Vetter, B. 2013. 'Multi-track Dispositions', *Philosophical Quarterly,* 63(251): 330-352

Vihvelin, K. 2004 'Free Will Demystified', *Philosophical Topics* 32(1/2): 427-450

Vihvelin, K. 2013. *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. Oxford University Press.

Warington, R. 1852. 'On the habits of the water-snail and stickleback', *Annals and Magazine of Natural History* 2(10): 273-280

Williams, N. 2011. 'Putting Powers Back on Multi-Track', *Philosophia*, 39(3): 581-598

Yarali, A., Niewalda, T., Chen, Y. C., Tanimoto H., & Duerrnagel, S. 2008. '"Pain Relief" Learning in Fruit Flies', *Animal Behavior* 76(4): 1173-1185